GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

OHN E. ANDERSON
University of Minnesota

HARLOTTE BÜHLER
Universität, Wien

VILLIAM H. BURNHAM
Clark University

YRIL BURT
University of London

EONARD CARMICHAEL
Tufts College

D. CLAPARÈDE
Université de Genève

EDMUND S. CONKLIN
Indiana University

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation,
Boston

LETA S. HOLLINGWORTH
Teachers College,
Columbia University

BUFORD JOHNSON
The Johns Hopkins University

AROLD E. JONES
University of California

RUMAN L. KELLEY
Harvard University

YOSHIHIDE KUBO
Hiroshima Normal College

K. S. LASHLEY
Harvard University

A. R. LURIA
Medico-biological Institute, Moscow

TOSHIO NOGAMI
Kyoto Imperial University

IVAN P. PAVLOV
Gosudarstvennii Institut
Eksperimentalnoi Meditsini,
Leningrad

HENRI PIÉRON
Université de Paris

GEORGE D. STODDARD
State University of
Iowa

CALVIN P. STONE
Stanford University

LEWIS M. TERMAN
Stanford University

GODFREY THOMSON
University of Edinburgh

E. L. THORNDIKE
Teachers College,
Columbia University

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

HELEN THOMPSON WOOLLEY
Teachers College,
Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 21

1 9 3 9

Copyright, 1939, by The Journal Press
Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

TABLE OF CONTENTS

No. 1

An experimental analysis of "level of aspiration" 3
By *Rosalind Gould*

No. 2

Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii 119
By Madorah E. Smith

No. 3

Domination and social integration in the behavior of kindergarten children and teachers 287
By Harold H. Anderson

The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli 387
By William Egleston Galt

No. 4

The social-sex development of children 461
By Elise Hatt Campbell

$7.00 per volume
Single numbers $3.00

QUARTERLY

February, 1939
Volume 21, No. 1

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota
Charlotte Bühler
Universität, Wien
William H. Burnham
Clark University
Cyril Burt
University of London
Leonard Carmichael
Tufts College
Ed. Claparède
Université de Genève
Edmund S. Conklin
Indiana University
Sante De Sanctis
R. Università di Roma
Arnold Gesell
Yale University
William Healy
Judge Baker Foundation, Boston
Leta S. Hollingworth
Teachers College, Columbia University
Buford Johnson
The Johns Hopkins University
Harold E. Jones
University of California
Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal College
K. S. Lashley
Harvard University
A. R. Luria
Medico-biological Institute, Moscow
Toshio Nogami
Kyoto Imperial University
Henri Piéron
Université de Paris
George D. Stoddard
State University of Iowa
Calvin P. Stone
Stanford University
Lewis M. Terman
Stanford University
Godfrey Thomson
University of Edinburgh
E. L. Thorndike
Teachers College, Columbia University
C. J. Warden
Columbia University
John B. Watson
New York City
Helen Thompson Woolley
Teachers College, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a limited contract, and may not be resold for a period of five years from date of publication.

Provincetown, Massachusetts
Copyright, 1939, by The Journal Press

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the Act of March 3, 1879.

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs,* 1939, **21,**
3-115.

AN EXPERIMENTAL ANALYSIS OF "LEVEL OF ASPIRATION"*

A dissertation submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in the Faculty of Philosophy,
Columbia University
1938

ROSALIND GOULD

Department of Psychology, Columbia University

Acknowledgments 5

I. The problem 7

II. The procedure 15

III. Statistical evaluation of generality 21

IV. Analysis of interview material 39

V. Analysis of interview material (continued) 67

VI. Discussion of conclusions 103

VII. Summary 109

References 115

---

*Received in the Editorial Office on October 20, 1938, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

ACKNOWLEDGMENTS

The writer desires to express her gratitude to Professor Gardner Murphy for his unfailing encouragement and aid, and to Dr. Max Hertzman and Mr. Isidore Chein for their discriminating criticism of the manuscript. The writer also wishes to thank Drs. O. Klineberg, D. W. Chapman, and J. Volkmann for their coöperation in permitting the administration of the various tests to their students.

Rosalind Gould

439 West 123rd Street
New Yorkk City

I. THE PROBLEM

The first study on *Anspruchsniveau,* translated somewhat inadequately as the "level of aspiration," grew out of the Gestalt-inspired experiments on both the effects of forced stopping of tasks before their completion, and on the nature of the psychological background preceding spontaneous breaking off and spontaneous resumption of a task. Such considerations led to the realization that feelings of pleasure or displeasure with the task, arising from the feelings of failure or success associated with the obtained performance, determined the subsequent behavior (of breaking off or resumption, e.g.) and the attitude of the subjects to that task or similar tasks. This approach to the problem revealed that the same objective performance score was interpreted differently by different individuals depending on their different standards. Thus feelings of failure and success were dependent almost entirely on the individual's standards of performance or his "goals" of performance in a given task and situation, rather than on the objective task-instructions. Hoppe (7) then attempted to study the reciprocal relationship between goals and feelings of failure and success, i.e., how the latter affected the height of the goals and how the height of the goals determined whether a given performance level would be interpreted as failure or success. It must be borne in mind, of course, that the experience of failure and success is not strictly bound to particular performances so that a given objective achievement is not psychologically the same experience at different times for the same individual, nor for different individuals at the same time. An individual who has achieved 80 after having obtained only 60's previously would be inclined to regard 80 as success, but after attaining 90 several times, or even once perhaps, 80 may be considered in the light of a failure. Since individuals, moreover, have different standards and background of experiences, the same objective performance score will have different psychological meanings for different persons. It also follows that the absence of a goal or "momentary level of striving" would make impossible the experience of success or failure in relation to a performance.

The height of the individual's goal or "aspiration level" was ascertained by Hoppe largely in qualitative terms, so that exactness and precision were well nigh impossible. One of his criteria was, e.g.,

> ". . . the appearance of success or failure experiences from which the direction of the aspiration level can be derived. . . . From success experiences one can derive that aspiration level lies below or is as high as the just achieved performance; from failure experiences that aspiration level lies higher" (7, p. 11-12).

Sometimes the subjects were more specific in their behavior and made such remarks as *"First I will try this figure,"* or *"I must at least achieve 100 once"* (7, p. 12). On the whole, such criteria are at best provocative and at worst conducive to misinterpretation. Studies on expressive movements and their interpretation have given us little basis for relying on such criteria.

What enabled Hoppe to study the dynamics of the problem with essential accuracy was both the fact that the actual performance scores were recorded in some of the tasks and that introspections were obtained at the close of the experiment. Important too was the feeling of freedom in the situation, since the subjects were not compelled to complete a task but could stop at any time they desired, as they could also return to any task they wished. A variety of problems was used, nine in all, ranging from a thought problem to a task of throwing rings on a moving target. Unfortunately some of them were really all-or-none tasks which defined "failure" and "success" in necessarily arbitrary, rather than in individual terms, and *all* of them defied quantitative treatment of the relationship between performance and aspiration level. Despite these drawbacks Hoppe observed certain general changes in aspiration level in all of his four main subjects, i.e., with success it was increased and with failure there was a tendency for it to be decreased. Hoppe presents these changes graphically, but since the different aspiration levels had no strict meaning (reality), the curves indicate relationships more clearly than is justified.

The disadvantage of using too few subjects, of too much dependence on impressionistic criteria, and of too rigid or at times too fluid an experimental structure, does not prevent this study from making a highly significant contribution. The demonstration of striking individual differences in definitions of and reactions to "success" and "failure," and the greater or lesser responsiveness of aspiration level to "success" and "failure" led the way to the realization that the momentary goal-strivings in relation to a given task were not only dependent on the goals set by the task-instructions, or the actual level of ability of the person. These responses were also related to the individual's deeper strivings which motivated his ambitions and behavior in everyday life.

Hausmann (6), recognizing that the reactions to success and failure provide a significant clue to the personality organization or disorganization of the individual, utilized the aspiration level technique "to evaluate personality traits" of 150 subjects, most of whom were psychoneurotics and pre-psychotics. His approach was largely qualitative despite his recording of the performance scores and the individual's estimate of future performance. He analyzed these scores in such terms as the speed with which estimates were adjusted upwards or downwards and the large or small steps in such adjustment. On the basis of these analyses he came to certain conclusions regarding such characteristics of the individual as "aggressiveness," "perseverance," "stubborness," and "instability," in one case even attempting to link this psychological instability to evidence of instability of the autonomic nervous system.

His procedure differed from Hoppe's in that an arbitrary number of trials had to be performed by the subject in a single task, dart throwing, and the usual experimental question, namely, *"What will you do next time?"* or *"What do you intend to do next time?"* was altered and made more complex. Hausmann's subjects, in effect, contract to obtain a given performance score. If they just make the bid or go above it they are given the same credit, but if their performance scores are below the bid, the number of points lacking is doubled and that sum subtracted from the bid. Thus the subject is penalized for going above or below his "contract" since he is given no credit for going above, and if he goes below, his recorded score, for all apparent purposes, is less than he actually achieved. This procedure places a premium on "accuracy" of estimation, and in effect attempts to eliminate differences in attitude towards accuracy of estimation as an uncontrolled variable. Such a procedure may, however, have the unfortunate effect of reducing individual differences in the expression of true aspiration level. I.e., the subject may actually have a higher aspiration level than the need of caution will permit him to express. It might be argued that under such conditions differences which are found between estimated performance and actual performance would be more significant as measures of aspiration level; i.e., discrepancy between performance and estimate, despite the repressive effect of need for accuracy, is a significant index of the difference between desired level of performance and the ability to perform at that level. It is most reasonable to suppose, however, that differences under these conditions are rather an index of the subject's poor judgment than of basic personality trends. The need for accuracy, in effect, we believe, eliminates the possibility of giving expression to one's aspiration level.[1]

This method tends to obscure individual definitions of failure and success which are necessary for adequate interpretation of the data, since it is apparently assumed that failure is inevitably experienced when the bid is not made, and that success must follow when the bid is just attained. This would overlook such factors as the number of points away from the estimate and the direction of the disparity between performance and estimate, i.e., whether performance was above or below, all of which, albeit differentially for different persons, might conceivably determine the nature of the failure or success experiences. In addition, the meaning of the bid-score to the individual must be determined before coming to any conclusions, since in cases where the estimated score is considered "poor" to begin with, achieving it might not yield an experience of success. To understand the relation between the quantitative and qualitative data one must first go back to a consideration of the nature and function of the estimates themselves. Hausmann's approach, essentially uncritical, avoiding the controversial issues involved in the interpretation of such estimates as indicative of aspiration level, is an example of some of the recent aspiration level studies where stock experimental variables, strongly reminiscent of the superficial praise and blame experiments, are introduced into a situation charged with unknown factors. It is unfortunate that the best known American study (1, 2) was accepted, at least for a while, so uncritically and completely that its assumptions and inferences were taken for established facts.

---

[1] It may also be well here to make the point that since aspiration level is as yet an inadequately explored concept in the case of normal subjects, the application of this concept to special problems such as the study of psychoneurotics who may present greater complexities in their psychic structure, is as yet premature.

Frank (1), impressed chiefly by Hoppe's findings of the relationship between aspiration level and the character of the individual, undertook to study the problem of the generality of this measure of one's strivings. He improved the technique, particularly from the point of view of ease of quantification, by presenting each task for a given number of trials, and recording the obtained differences between performance and aspiration. Before each trial, after the preceding performance score had been announced, the subject was asked simply, *"What do you intend to do next time?"* Aspiration level was then defined in strictly operational terms as "the level of future performance in a familiar task which an individual, knowing his level of past performance in that task, explicitly undertakes to reach" (1, p. 119). It is clear that the estimates of future performance cannot be taken as a measure of the level of the individual's strivings. "Level" in itself implies a dynamic *relationship.* Consequently the only way of evaluating the height of this explicit estimate is in terms of the relationship between estimate and performance, i.e., the extent of the difference between them. Although by his definition, aspiration level becomes synonymous with the explicit estimate of future performance, Frank nevertheless regards the measure of the relationship between performance and estimate as but one "aspect" of the "level of aspiration."[2] It would appear that the only measure of the *height* of aspiration level as herein defined by Frank would be the extent and direction of the difference score. (This measure will hereafter be referred to as the *average difference* score, since it is obtained by taking the difference between the average, or in some cases the median of the performances, and the average of the estimates in the same task.) This in no way denies, however, that a psychologically rather than an operationally defined aspirational level would be different from a difference score, nor do we overlook the fact that other problems, such as variability, are involved in the relationship between estimates and performance scores.

It is this average difference score, rather than the aspiration level which Frank concludes is "a relatively permanent characteristic of the personality . . . which permanence can be demonstrated regardless of the type of ability which the task requires" (1, p. 123). The difference score is said to be dependent primarily on the relative strength of three needs: "(*a*) The need to make the level of aspiration approximate the level of future performance as closely as possible; (*b*) the need to keep the level of aspiration high regardless of the level of performance; and (*c*) the need to avoid failure" (2, p. 285). Should these conclusions be valid, the implications for future psychological research in personality development are many. E.g., under what bio-environmental circumstances does a particular need develop? How clearly does it begin to function? Is one need definitely more important for adjustment than another? Is it better, that is, to overestimate than underestimate one's level of future performance? If one is more satisfactory, then it is essential to know the possibilities of substituting for an inadequate method of reacting.[3] Or, is it true here, as in other descriptions of types, that most persons utilize all three "needs," which one depending primarily on the situation, and that "pure" types in the sense of only one driving need are rare? Further, what are such levels of aspiration related to? Would we be able to find definite relationships with culture, religion, socio-economic status, age, political principles, and so on?

---

[2] Since Hoppe points out that aspiration level is a function of the "need to keep aspiration level as high as possible," i.e., the need to achieve success, and the need to avoid failure, Frank, by attributing the dependence of the difference score to these same needs, implicitly assumes an essential identity between aspiration level and the difference score.

Before such questions can be answered one must first determine the nature of the difference score and its operation under adequately controlled conditions. It is important at this point to analyze Frank's data in detail, since the criticisms directed against his experimental situation define the writer's problem in the clearest way possible. In his experiment there were three groups of 12 subjects each, and three tasks (printing: *P*; spatial relations: *S*; and quoits: *Q*), which were divided so that each group had two of the tasks. Thus Group I had Tasks *P* and *S*, Group II, *S* and *Q*, and Group III, *Q* and *P*. Each task was given twice, in two sessions with about a week intervening between each session. Throughout the statistical analysis of his data Frank employs what we have referred to as the average difference score, which in his study means the difference between the average of the estimates and the median of the performance in a single task. Fifteen trials were given in each task. Discarding the *first few trials as practice, the median of the performance scores on* trials 5-14 was subtracted from the average of the estimates in *trials 6-14.*

---

[3] Updegraff and Keister have demonstrated that within a given period of time unsatisfactory reactions of pre-school children to failure, e.g., rationalizing, asking for help continually, lack of perseverance, and so on, can be changed to satisfactory habits of meeting obstacles. It is not yet known, however, how long the induced habits are maintained without further training (18).

The first step is to establish the reliability of the difference score which he does by correlating the difference scores for the same task obtained in the two sessions, i.e., by the test-retest method. Table 1 presents the results. It is at once evident that the choice of quoits was unfortunate. The marked difference between what should be corresponding correlations for the two sessions of quoits indicates its low reliability. For the other two tasks, while the correlations do not vary much from each other and are significantly greater than zero the fact that each is based on only 12 subjects makes them far from conclusive. Nevertheless, Frank concludes that the difference score is a relatively stable characteristic of the individual. The possible relationship between any given level of performance and its *corresponding aspiration level is not considered, nor is the correlation* of average difference scores alone, conducive to providing a clear idea of the dynamic interrelationship between performance and aspiration. We do know, for example, whether changes in size of the difference score occurring between test and retest are largely a function of changes in performance level without adequate adjustment of the aspiration level, or changes in aspiration level with performance level more or less constant, or changes in both performance score and aspiration level?

TABLE 1

	Printing	Spatial relations	Quoits
Group I	.72	.57	
Group II	.60		.26
Group III		.75	.63

TABLE 2

	Group I		II		III	
(*A*)	$(P_1+P_2)/2$, $(S_1+S_2)/2$	.62	$(Q_1+Q_2)/2$, $(P_1+P_2)/2$	.62	$(Q_1+Q_2)/2$, $(S_1+S_2)/2$	.37
(*B*)	P_1, S_1	.50	Q_1, P_1	.70	Q_1, S_1	.45
(*C*)	P_2, S_2	.59	Q_2, P_2	.01	Q_2, S_2	—.08
(*D*)	P_1, S_2	.54	Q_1, P_2	.45	Q_1, S_2	.26
(*E*)	P_2, S_1	.65	Q_2, P_1	.06	Q_2, S_1	—.06

Utilizing the difference score again, Frank studied the generality of this measure by correlating the difference scores of two tasks, each obtained over two sessions, with each other. Table 2 presents the obtained correlations. The rank order correlations between both sessions of *P* and *S* is .62, between *Q* and *P*, .62, between *Q* and *S*, .37. The finer analysis of the first and second sessions of each task reveals that wherever the second session of *Q* enters in, the correlation becomes about zero. Frank disregards the latter measures and concludes on the basis of the lumped data, that the correlations .62, .62, and .37 reveal "a single trait which manifests itself both *in* tasks *differing* in performance scale and in tasks requiring different types of ability" (1, p. 124). The only explanation of the fact that Group III has a correlation of .37 is that tasks *Q* and *S* differ in both the common elements of "speed" and "motor ability" which characterize the combinations of *P* and *S*, and *P* and *Q*, respectively. Two other hypotheses are just as likely: (*a*) the drop in correlation may be due solely to the unreliability of the quoits task, or (*b*) all 3 correlations merely indicate the degree of homogeneity of requisite ability or level of difficulty among the tasks themselves, rather than the existence of a single trait.

Frank's conclusions then, regarding the generality of the difference score, his measure of aspiration level, loses some of its significance in view of the weaknesses of the experimental structure, namely, (*a*) failure to consider either the possible effect of level of performance on the range of aspiration-estimates, or (*b*) the possible distorting influence of the relationship of performance ability in different tasks to one another, and (*c*) the use of a small number of both tasks and persons. If any conclusions are to be drawn concerning the generality of any measure, particularly in a correlational study, it is necessary to present a large number of tasks to an adequate sampling of individuals.

Since much research on aspiration level hinges on this point, the main problem of the present study is to determine the validity of the assumption that the difference score is a relatively stable characteristic of the individual. The extent to which this assumption is valid can be ascertained by the consistency with which both the direction and amount of difference between performance scores and estimates of future performance is found in a given individual, throughout a variety of unrelated materials. It was with these factors in mind that the present experimental plan was drawn up.

II. THE PROCEDURE

A. SUBJECTS

The subjects were Columbia male undergraduates, mainly sophomores who were enrolled in the introductory psychology course. They constitute a fairly homogeneous group with respect to age and educational status, factors whose influence on aspiration level has not yet been determined. The main laboratory investigation involved 82 subjects.[4]

B. TASKS

The choice of tasks to use was determined on the basis of several criteria: (*a*) To have as varied a sample of materials as possible, (*b*) to have the tasks uncorrelated in performance with each other, (*c*) to have tasks with high reliabilities, and (*d*) to have tasks vary in type of performance curves obtained, i.e., to have tasks where learning definitely occurs and tasks where performances remain on essentially the same level.

The following six tasks were found, on the basis of preliminary results, to meet the above requirements.

(*a*). *Synonyms* test. Nine lists of 33 words each were devised, one of the lists being used for practice. Each list was treated as a separate trial. The subject was asked to write after each stimulus word, or after as many words on each list as he could, a word meaning the same or nearly the same. For four of the lists used in the experiment proper, the subject was instructed to write synonyms beginning with the letter *A*; for the other four they were to begin with *B*. (E.g., stream—brook; graze—browse; help—aid or assistance, etc.) The *A* and *B* lists were alternated. Two minutes were allowed for the completion of as much of each list as possible. After each trial, the sheet was scored immediately, the subject told his score, and asked, "*What will you do next time?*"

It is important that the lists be of approximately equal difficulty. Originally the stimulus words and as many of their associated answers as possible, were roughly equated on the basis of frequency

---

[4] In addition to this investigation, two other problems were studied concurrently: (*a*) The problem of the relationship of the difference scores to other more or less accepted measures of personality (5), and (*b*) the problem of the relationship between estimates obtained in a life-situation and those obtained in this laboratory situation (4).

III. STATISTICAL EVALUATION OF GENERALITY

A. Reliabilities of Tests Used

All reliabilities reported in Table 4 are split-half reliability coefficients corrected by the Spearman-Brown formula. The reliabilities of the performance scores range from .89 to .99. The reliabilities of the corresponding "aspiration" estimates range from .95 to .99.

TABLE 4

Reliabilities of Performance and Estimates and the Correlations between Performance and Estimates

Tasks	Performance	Estimates	Correl. between perf. and estimates
1. Words	.90	.95	—.03
2. Stead.	.94	.96	.11
3. Add.	.99	.99	.10
4. *S-D*	.97	.96	.06
5. Canc.	.92	.97	—.13
6. Target	.89	.97	.08

B. The Correlations

Since the problem of adequate sampling is a crucial one, it will be important first to establish the stabilization of the average difference intercorrelations.

1. *Stability of the Average Difference Intercorrelations.* When rank-difference correlations of the difference scores of the first

TABLE 5

Comparison of Average Difference Intercorrelations Using Two Correlational Techniques

Tasks	1	2	3	4	5	6	Mdn A^1	Mdn A^2
1. Words		.39	.31	.43	.40	.57	.40	.35
2. Stead.	.41		.43	.29	.33	.21	.33	.08
3. Add.	.28	.43		.60	.36	.36	.36	.43
4. *S-D*	.37	.08	.57		.47	.49	.47	.37
5. Canc.	.26	.05	.12	.32		.40	.40	.26
6. Target	.35	.04	.43	.45	.52		.40	.43
						Medians	.40	.35

[1] Median *A* refers to the correlations above the line which were obtained by the rank-difference method. $N = 33$ (cases 13 to 45).

[2] Median *B* refers to the correlations below the line which were obtained by the product-moment method. $N = 30$ (cases 13 to 42).

33 complete cases (with subjects 1-12 omitted) were calculated, a median of .40 was obtained. In the final analysis, using the product-moment method for the 82 cases, a median of .29 was obtained. This difference in the size of the medians may be a function of the correlational techniques or may be due simply to the fact that the correlations had not become stable at 33 cases. In Table 5 we find a wider range in size of correlations when the product-moment method is used, although the median of the two groups is practically the same. What is perhaps most striking is the differential effect of the two methods on different tasks (cf. the steadiness task, e.g.), which also clearly indicates the complete unreliability of correlations obtained on few subjects.

The experiment is of such a nature that obtaining data for a single subject required the expenditure of a considerable amount of time and effort. It was therefore important to ascertain whether the rather restricted sampling of 82 cases was sufficient to obtain stable correlation coefficients. Intercorrelations among the differ-

TABLE 6

Tasks:	II	III	IV	V	VI
I. *A*[2]	.07	.02	.04	.05	.08
B	.04	.01	.02	.09	.03
C	.02	.02	—	.01	.01
II. *A*		.01	.13	.34	.14
B		.13	.03	.08	.09
C		.05	—	.06	.05
III. *A*			.22	.11	.27
B			.09	.17	.11
C			—	.00	.01
IV. *A*				.06	.06
B				.04	.04
C				—	—
V. *A*					.11
B					.05
C					.02

[1] The 12 individuals who represent the difference between 70 and 82 cases were not measured on variable IV.

[2] *A* represents the absolute difference between corresponding correlations based on 30 and 50 cases. *B* represents the absolute difference between corresponding correlations based on 50 and 70 cases. *C* represents the absolute difference between corresponding correlations based on 70 and 82 cases.

ence scores were therefore obtained for varying numbers of cases. Samplings of 30, 50, 70, and the full 82 were used. It was felt that we could consider that our correlations had stabilized (in view of no known true correlation with which to compare our results) when differences between successive samples were negligible. Accordingly Table 6 was devised. In this table, we present the absolute differences between corresponding correlations of samplings of 30 and 50 which we call the *A* difference, and similarly, differences between corresponding correlations for 50 and 70 cases (*B*) and for 70 and 82 cases (*C*). For the *A* differences the range is .01 to .34 with a median absolute difference of .08. For the *B*, there is a narrowing to a range from .01 to .17, with a median of .05. For the *C* set, the differences have become very small, ranging from .01 to .06 with a median of .02. Table 5 indicates the actual size of the intercorrelations for the first 30 complete cases (1-12 omitted); Table 7 for the first 50 and 70; and Table 8 for the complete group of 82 cases.

TABLE 7

AVERAGE DIFFERENCE INTERCORRELATIONS

Tasks	1	2	3	4	5	6	Mdn *A*[1]	Mdn *B*[2]
1. Words		.38	.27	.35	.30	.24	.30	.27
2. Stead.	.34		.31	.24	.31	.09	.31	.34
3. Add.	.26	.44		.44	.40	.27	.31	.26
4. *S-D*	.33	.21	.35		.34	.35	.35	.35
5. Canc.	.21	.39	.23	.38		.36	.34	.38
6. Target	.27	.18	.16	.39	.41		.27	.27
						Medians	.31	.33

[1] Median *A* refers to the correlations above the line where $N = 70$.
[2] Median *B* refers to the correlations below the line where $N = 50$.

TABLE 8

AVERAGE DIFFERENCE INTERCORRELATIONS WHERE $N = 82$—THE COMPLETE NUMBER OF CASES

Tasks	2	3	4	5	6	Median
1. Words	.36	.25	.35	.29	.23	.29
2. Stead.		.26	.24	.25	.04	.25
3. Add.			.44	.40	.28	.28
4. *S-D*				.34	.35	.35
5. Canc.					.34	.34
6. Target						.28
						Median = .29

We feel that the small differences existing between the corresponding correlations obtained from the samplings of 70 and 82 cases justifies our treating the correlations based on the 82 cases as very representative of the interrelationships among the difference scores in this study.

2. *Correlations Between Average Difference Scores and Performance Scores.* Before one can evaluate the above correlations it is necessary to investigate the possible effect of level of performance on the range of estimates and the possible distorting influence of similar task materials which would act to raise spuriously the average difference intercorrelations. Table 4 presents the correlations obtained between difference and performance scores. All six tasks, it is seen, have insignificant correlations, the largest correlation being —.13 for cancellation, with a *PE* of ±.07. This means that any difference score, regardless of size, is equally probable at any level of performance. The introduction of any special statistical treatment that would counteract distortions produced by significant correlations between performance and aspiration, becomes unnecessary. Even in this connection the necessity of adequate sampling is apparent, since an evaluation of the results obtained on the first 33 cases revealed a sizable correlation between the difference score and performance score in two of the tasks, addition and steadiness. Any task having a readily attainable maximum score would have a constraining effect on aspiration level and consequently on the difference score. E.g., a few students obtained a zero number of contacts at different times on the steadiness test. If any appreciable number had done so, or if any subject had done so consistently, the correlation between performance score and average difference score would be a significant positive one. Moreover, it is not unlikely, even under these conditions where in four tasks the maximum performance score is unknown, and in two tasks (steadiness and target) though known it is rarely attained, that the level of performance would have had more of a constraining influence on the range of aspiration level if the subjects had been aware of the "goodness" of their performance scores in relation to other subjects. Although most of the subjects asked in one way or another, *"What is the average?"* none was given any indication of his relative standing and since each subject kept his promise not to reveal the nature

of the experiment, such information was not available to any participant.[6]

3. *Intercorrelations of Performance Scores.* Table 9 presents

TABLE 9

AVERAGE PERFORMANCE INTERCORRELATIONS

Tasks	2	3	4	5	6	Median
1. Words	—.14	.19	.22	.10	—.11	.10
2. Stead.		—.36	.05	.08	—.60	—.14
3. Add.			.36	.27	.04	.19
4. *S-D*				.27	.17	.22
5. Canc.					.12	.12
6. Target						.04
					Median =	.10

Where variable 4 is involved $N = 70$ instead of 82 cases.

these intercorrelations. Although preliminary investigation indicated approximately zero performance intercorrelations, in the final analysis we find that 4 of the 15 correlations are above .20, the highest being .36. There are also two high negative correlations of —.36 and —.60, in both cases involving the steadiness test. We find that those average difference intercorrelations above the median of .29 also have higher performance intercorrelations, .18, than those average difference intercorrelations below the median, which have an average performance intercorrelation of —.08. If, however, we compare the median of the average difference intercorrelations of the variables showing relatively high performance intercorrelations, with the median of the average difference intercorrelations of the variables showing relatively low performance intercorrelations, which are .35 and .25 respectively, it is clear that the effect of the undesired homogeneity of material in this set-up, on the size of the difference intercorrelations, is slight. The trend, however, indicates that interrelated materials will make for higher average difference intercorrelations than would otherwise be the case.

4. *Groupings.* In order to examine the influence of material presented in same and different sessions, three groups were made as described in Table 3, and correlations were obtained for each of

---

[6] A study is now being planned by the author where this factor of "goodness" is introduced as an experimental variable.

TABLE 10

AVERAGE DIFFERENCE INTERCORRELATIONS OF THE THREE GROUPS SEPARATELY

Group I ($N = 28$)

No.	1	5	6	2	3	4
1	—	.19	.15	.47	.34	.49
5			.36	—.14	.29	.20
6				—.13	.22	.13
2					.52	.60
3						.49
4						—

Group II ($N = 28$)

No.	1	2	6	3	4	5
1	—	.52	.13	.14	.31	.23
2			.21	—.02	.15	.38
6				.26	.44	.30
3					.46	.61
4						.59
5						—

Group III ($N = 26$)

No.	1	3	6	2	4	5
1	—	.45	.46	—.08	.31	.40
3			.37	.37	.37	.46
6				.17	.52	.40
2					.19	.61
4						.45
5						—

the groups separately. Table 10 reveals again the importance of sampling of both tasks and persons. E.g., correlations between words and target, in all three cases appearing in the same session, vary from .13 to .46, correlations between steadiness and symbol-digit vary from .60 in Group I to .19 in Group III. On the other hand correlations between addition and symbol-digit, .49 and .46, appearing in the same session in Groups I and II, and between cancellation and symbol-digit, .45 and .59 in Groups II and III, show relatively slight variations. A clearer picture of these differences appears in Table 11, where despite differences in the composition of the groups,, one sees clearly that on the whole higher correlations obtain between tasks appearing in the same session, than between those appearing in different sessions. The median for all three groups of tasks appearing in the same session is .45 as com-

TABLE II

COMPARISON OF AVERAGE DIFFERENCE INTERCORRELATIONS OF TASKS APPEARING IN SAME AND DIFFERENT SESSIONS

Group I

Same		Different	
Tasks		Tasks	
1, 5	.19	1, 2	.47
1, 6	.15	1, 3	.34
5, 6	.36	1, 4	.49
2, 3	.52	2, 5	—.14
2, 4	.60	3, 5	.29
3, 4	.49	4, 5	.20
		2, 6	—.13
		3, 6	.22
		4, 6	.13
Median = .425		Median = .22	

Group II

Same		Different	
Tasks		Tasks	
1, 2	.52	1, 3	.14
1, 6	.13	1, 4	.31
2, 6	.21	1, 5	.23
3, 4	.46	2, 3	—.02
3, 5	.61	2, 4	.15
4, 5	.59	2, 5	.38
		3, 6	.26
		4, 6	.44
		5, 6	.30
Median = .49		Median = .26	

Group III

Tasks		Tasks	
	Same		Different
1, 3	.45	1, 2	—.08
1, 6	.46	1, 4	.31
3, 6	.37	1, 5	.40
2, 4	.19	2, 3	.37
2, 5	.61	3, 4	.37
4, 5	.45	3, 5	.46
		2, 6	.17
		4, 6	.52
		5, 6	.40
Median = .45		Median = .37	

NOTE: Median of tasks appearing in same session for total group = .455. Median of tasks appearing in different session for total group = .30.

pared with .30 for tasks taken at different times. That these medians represent a trend and not a rule is also clear from this table. E.g., the correlation between cancellation and target in the same session is .36 in Group I, the correlation between these two tasks appearing in different sessions is .30 for Group II and .40 for Group III. The range of correlations is tremendous when dealing with a small number of cases, so that such disparities in size of correlations as appear, for example, between symbol-digit and target, which tasks always appeared in different sessions (.13, .44, and .52 for the three groups) should not be taken too seriously.

5. *The Average Difference Intercorrelations.*

a. The influence of interrelated materials on the size of average difference intercorrelations. It will be recalled that Table 8 presents the average difference intercorrelations of the six tasks, revealing a range from .04 between steadiness and target to .44 between symbol-digit and addition. In the light of our discussion of the fact that interrelated materials will act to raise spuriously average difference intercorrelations, it is significant that the lowest difference intercorrelation is obtained by the task having the highest negative performance intercorrelation, namely —.60 between steadiness and target, and the highest difference intercorrelations are obtained by those tasks having the highest positive performance intercorrelations. Addition and symbol-digit, addition and cancellation, and symbol-digit and cancellation, having average difference intercorrelations of .44, .40, and .34 respectively, have performance intercorrelations of .36, .27, and .27 respectively. That the factor of homogeneity of material is not the main determinant of the size of the difference intercorrelation, at least that it does not function in any simple and direct manner, is revealed by the fact that average difference intercorrelations of .35 (words and symbol-digit), .35 (symbol-digit and target), and .34 (cancellation and target), have associated performance intercorrelations of .19, .17, and .12 respectively. There is no doubt, however, that tasks related in performance to any marked degree will yield average difference intercorrelations that are spuriously high. Since there is no record of the performance intercorrelations of the tasks employed by Frank, we have no way of judging whether this factor alone, or the combination of similar material and select sampling, accounts for the marked difference in our results, and consequently in the conclusions drawn from the statistical data.

b. The extent of generality of the average difference score. We obtain a median of .29 based on 15 intercorrelations, in contrast to correlations of .62, .62, and .37 obtained by Frank. He naturally concluded that the difference score demonstrated generality. In this study, since all but two of the 15 difference intercorrelations are reliably greater than zero,[7] and since both the tasks and the estimates have split-half reliabilities of .89 or above, we can conclude that there is better than chance relationship existing among the average difference scores. What then can we say about the generality of this measure?

In terms of our experimental situation where the materials are at most slightly intercorrelated, the group responds differently to different materials. The predictive value of an average difference score obtained in one task for an unrelated task is about 8 per cent above chance. This fact suggests the problem: what is the relation between "variability" and the size of the average difference scores? By "variability" here we do not mean the variations in the differences between performance and estimate occuring in a single task,[8] but the variability in the size of average difference scores obtained by a given individual from one task to another. As indicated above, different materials. The predictive value of an average difference score into comparable distributions, and the measure of variability was the sum of the differences of the six transmuted scores from the median of these six difference scores. Taking the group of 82 subjects as a whole, we find a barely significant correlation of .33 with a *PE* of ±.07 between this measure of variability and the median average difference score. This would mean that if people differ with respect to generality of aspiration level, this difference is not primarily a function of the level of aspiration. That is, high variability, which indicates a different organization of motives from that which is present when low variability is obtained, bears little relation to the height of the aspiration level, although the correlation does indicate a slight tendency for those having relatively low median difference scores to have low variability, in other words to respond more to the situation than to the specific task.[9]

---

[7] This estimate is based on accepting a ratio which indicates 99 or more chances in 100 of a true difference as representing a true correlation.

[8] This problem will be discussed in a separate paper to be published shortly.

C. The Average Difference Score

1. *High (H), Medium (M), and Low (L) Average Difference Score Groups.* In order to give psychological meaning to the difference scores, it would be desirable to study separately the three groups differentiated quantitatively on the basis of size of average difference score, to see whether there are differential underlying patterns of motivation. Unfortunately we have no reason to choose one task as our criterion for division rather than another, and the low inter-task correlation points to wide fluctuations in difference scores for many individuals. Table 12 presents the obtained reduced median difference scores of the 82 subjects. Our point of reference at this stage should be the quantitative results, but these results

TABLE 12

The Median Reduced Scores of the High (*H*), Low (*L*) and Medium (*M*) Average Difference Score Groups

High		Low		Medium			
Subject	Md.	Subject	Md.	Subject	Md.	Subject	Md.
7	55	1	38	2	51	49	44
13	56	9	35	3	52	50	51.5
15	56.5	14	34.5	4	51	51	44
17	58.5	16	38.5	5	48	55	50.5
18	59	22	37.5	6	44	57	54.5
19	59	31	42	8	54	58	49
23	64.5	42	36.5	10	46	59	50.5
24	61	47	39.5	11	43	61	47.5
25	58	48	37.5	12	53	62	49
28	60.5	52	35.5	20	54	66	52
29	55	53	33.5	21	48	67	52.5
33	62	54	34.5	26	50.5	68	43.5
35	59.5	56	41.5	27	49	69	51.5
38	59.5	60	41.5	30	43.5	71	46.5
41	62	63	36	32	47	73	45
43	57	64	37.5	34	51	74	53
44	55	70	31.5	36	45	75	48.5
45	62.5	77	37.5	37	42.5	80	48.5
65	62.5	78	39.5	39	50	81	46
72	68	79	32	40	54.5	82	43
76	59			46	51		

---

[9] In the intra-test analysis to follow this study, consideration will be given to the question of the differential variability of performance and aspiration in "high" and "low" variability groups.

do not represent the high degree of uniformity that we would prefer for the procedure that has been suggested, i.e., utilizing the median of the six transmuted difference scores as a single measure for differentiating individuals according to height of average difference scores. If we remember, however, that the aim of the segregation is to get some start for the analysis of patterns, this point is not an overwhelming objection. It means merely that there is a greater likelihood of overlap in the three groups if the division is based on all six tasks. The three groups, i.e., high, medium, and low average difference score groups, have actually been selected on the basis of a median reduced score. The analysis below shows very little overlap between scores on individual tasks among the members of the two extreme groups. Separation on the basis of quantitative results can be accomplished in this way to the extent permitted by the nature of the results. In these reduced scores, 50 was the mean. In the H group, out of 125 difference scores we find 27 scores of 50 or below. Out of 124 difference scores in the *L* group we find 108 scores of 50 or below.

Several considerations must be kept in mind in the analysis of the motives of the individuals in the 3 groups. A given motive may be expressed in different quantitative terms by different individuals, while different motives may find expression through the same quantitative means. Thus the meaning of any given difference score will depend on its particular source of motivation. If this is true then the division according to the difference score may be of no value. But there does remain the possibility of different orders of difference scores being symptomatic of some kind of organization to a greater extent than others are. The division becomes valuable to the extent to which it shows us how much or perhaps how little the difference scores can reveal of the basic patterns underlying the responses.

2. *Direction of Difference Scores.* Before a final evaluation of the problem of generality can be made, we must consider not only the consistency of amount of difference between performance and estimate but the consistency of direction of the average difference scores. "Direction" refers to the gross relationship between average performance and average estimate. I.e., regardless of the amount of difference between them, the difference score is positive if the average estimate is higher than the average performance, and if the average estimate

is lower, the difference score is negative. The question immediately arises, is the direction a discriminating aspect of the difference score? Out of 70 cases having 6 average difference scores, we find 65 per cent have 5 or 6 plus difference scores, 23 per cent have 3 or 4, and 10 per cent have 2 plus and 4 minus difference scores. One subject, or 1 per cent, has 1 plus and 5 minus scores, while only one has all 6 minus difference scores. This distribution of average difference scores could occur by chance not more than one in a million times. Factors are operating in this situation and in this group, making for a more or less uniform tendency to estimate scores higher than the achieved scores. What are these factors? A complete answer to this question involves the interpretation of the function of the aspiration-estimate, which interpretation is based chiefly on the interview material and will be presented in the next section. What appears to be the immediate cause of the essentially non-discriminating character of the direction of the difference scores, namely cultural factors, will be discussed at this point.

Anthropologists, sociologists, and psychologists have long recognized the importance of "culture" in its selective stress on certain types of behavior-reactions. The preponderance of plus difference scores[10] seems to us to be essentially a cultural manifestation rather than an inevitable reaction arising from the situation,[11] namely, a manifestation of the well-known cultural optimism or emotional "weighting of the future" which many sociologists have commented upon. Compare *Middletown in Transition* (11) e.g., where for a long time people were reluctant to accept the *fact* of a depression and concentrated all wishing, planning, and thinking on "tomorrow."

With practically no exception, in this situation the subjects, with differential intensity of desire in different individuals of course, always wanted to do better than they had just done. On the basis of the higher estimates one would conclude that they necessarily also *expected* to do better the next time. When we realize, however, the force of a cultural pattern, we understand that not only does weighting of the future connote that next time *will* be better, but it also has a fairly insistent implication of *obliging* one to be better next time.[12] The experimental question, *"What will you do next time?"* for many subjects meant, *"How much better will you do next time?"* So clearly was this obligation felt that some subjects, who were not sure of doing better, became quite uncomfortable, and in reply to the question anxiously stated, *"I doubt if it will improve much."* We thus had provision for a situation where, whether one actually felt he would do better or whether one had doubts about it, it was somewhat imperative to keep up a "front" of expecting to improve. Such an attitude is closely related to the need of appearing self-confident in our society, and it is not surprising that one of our subjects with unusual frankness and insight into some aspects of his behavior should admit that he interpreted the experimental question literally, i.e., *"How many problems will I do. . . .?"* but that *"Of course there's a feeling that you don't want to say you're going to do worse. . . . I wouldn't let it come into my mind. . . . I don't think there's any point in letting the other fellow know you're not confident. . . . I try to act confident in everything although some things I'm not sure of my capabilities in."* He is an extreme product of our mores, certainly, but we are in no way suggesting universality in the internalization or intensity of this attitude towards the future, merely that it is one of the important factors making for a much greater than chance occurrence of plus difference scores. Universality of response is not a requisite criterion of the presence of cultural determination. Few have been able to ignore individual differences within any cultural area, so that the above interpretation is in no way weakened by the absence of this attitude in some subjects. Finally, since there is cultural sanction in expecting to be or do better, even if one does not, there still remains the soothing fact that one *tried* and just trying to get ahead, though unattended by immediate success, is respected in this culture. That is to say, we have here an example of a culturally sanctioned rationalization of failure, which can be utilized readily when achievement falls below estimation of achievement.

---

[10] Frank's analysis of this phenomenon is discussed in the final chapter.

[11] It would be extremely valuable then, to study the nature of difference scores obtained in cultures differing widely from our own, e.g., the Zuni Indians, the Arapesh, and for a caricature of our society, the Kwakautles.

Another example of culturally determined attitudes in this situation is observed in the differential response to tasks labelled "intellectual" as opposed to those called "motor" or "athletic." Apparently to do well in athletics, although desirable, is not valued by this group (with striking exceptions) as highly as achievement in intellectual tasks. (Football heroes, according to tradition, should view the situation differently, but no representative of this group volunteered for the experiment.) The factors that operate most directly and differentially on these two types of materials are the still persistent undergraduate mores which brand an individual a "sissy" if he overtly expects to get high grades, and a "sissy" if he overtly expects to lose in a physical match. This highly discriminating attitude, however, is much weaker today than it has been for a long time. The depression and increased competition for jobs has presumably sobered the present-day college student, and no doubt the breaking down of the stereotype conceiving the "*A*" student as necessarily puny, "unmanly," has contributed to the change. However, the accompanying attitude regarding the "correct" behavior with respect to physical matches and extending to include all "athletic" endeavors, appears to be still strong. This means that an individual who wants to do a great deal better in the "motor" tasks can feel freer to estimate a score closer to his goal-desires than he might otherwise, since the penalty for "failing," i.e., going below his estimate, is not so great, inasmuch as the performance scores are not as important as in the "intellectual" tasks. On the other hand, as with the operation of the cultural factor discussed above, the demand of adhering to the custom may force some to act in a manner quite incompatible with their feelings, i.e., some may feel compelled to exhibit a bravado towards future scores in "athletics" which they do not feel, even estimating a score higher than they actually desire, much less expect to reach. The size of the mean average difference score in the different tasks reveals the trend of these attitudes in operation. The average of the three "intellectual" tasks, words, addition, and symbol-digit, is +1.0, in contrast to the average of +3.0 for the mean of the average difference scores in the three "motor" tasks. That the mean average difference score in the target is +5.3 is no cause for surprise, when we realize that not only is this task subject to the influence of the attitudes discussed, but it is also most often taken as a "game." The effect of a "play attitude" will be discussed in the final section.

---

[12] "It is important to note the strains which current cultural demands for dominance and aggression create in the individual personality. . . . False faces that the culture forces men to wear. Everywhere one is confronted by the demand that one . . . act 'like a man,' hide one's emotions, talk or appear 'successful,' . . . be 'sure of oneself' and so on indefinitely through the stereotypes of being 'regular'" (11, p. 427).

D. Summary and Discussion

The low intercorrelations of the six tasks do not substantiate the claim of consistency and generality of the average difference score when the latter is studied with functionally unrelated materials. It has been noted that despite the influence of cultural factors, which though operating differentially on different materials, at the same time produce a spurious appearance of uniformity of response, despite the raising of difference intercorrelations when performances in different tasks are positively correlated, and despite the effect of increase in size of correlation of tasks appearing in the same session as opposed to their appearing in different sessions, the median of the 15 average difference intercorrelations is .29. Though more than a chance relationship among these scores is hereby indicated, a correlation of .29 does not suggest the existence of a general factor. Is it possible that certain factors are operating here in a way to distort an underlying generality of response to the situation?

It is possible, in so unstructured a situation, for an individual who would ordinarily turn to one kind of response to be so completely lost that there is a rapid disintegration of his familiar reaction patterns. Further, temporary sets towards goal-ends may be established at the beginning of a task, so that the first few performance scores which may be atypical, nevertheless, in a few trials become the primary rather than the secondary determiners of the estimates and accompanying attitudes. That complete psychological chaos never occurs, however, and that some organizing principle still operates, is indicated by the very high reliabilities of the estimates. If there is an organizing principle operating within a task, why does it not operate in different tasks appearing in the same objective situation? The evidence for it not operating is of course, purely inferential, deduced from the fact that statistical analysis fails to reveal a "single trait" (1, p. 124). It may be urged that these results are artifacts of the situation since organized personality structures, such as we believe we are dealing with, are not readily viewed as self-inconsistent.

These results lose their character of self-inconsistency, become less enigmatic, when we realize the nature of this "personality" measure, the difference score between performance and "aspiration." We are simply dealing with an end-process, a complex overt behavior manifestation which has its roots and receives sustenance from sources largely below the surface. It is psychologically not implausible that quite different inner motivations will, under certain circumstances, make for apparently similar overt behavior responses, and as a corrolary, similar inner motivations and needs will produce apparently opposite behavior reactions. The same end may be achieved in different ways and from different starting points. In a study of the "sympathetic" and "aggressive" behavior of children, L. B. Murphy found that aggression is manifested as a response to an underlying feeling of self-confidence, and also arises in response to deep feelings of insecurity and anxiety (14). We have already observed how certain cultural demands may act to produce the same quantitative end in different individuals, although the underlying reaction to these demands may be completely opposed. Further corroboration of this point is obtained from the interviews. The difficulty of analyzing and interpreting average difference scores as a measure of aspiration level is a function of this fact—that one is dealing with the end-results of a process.

The individual simply is capable of responding in a number of ways, although the mainsprings of the behaviors may be the same. His aspiration level, e.g., "to be on top" or "to get 100 per cent" may not change necessarily, yet if asked to give an explicit estimate of his level of future performance, which means *committing himself*, under certain circumstances he may estimate low, and under others high. These data imply that in situations which seem to offer threat of potential failure, a number of reaction-patterns all functioning to avoid this threat are open to an individual, but one must not overlook the fact that not all such situations, however objectively similar, offer the same intensity of threat of failure. Given situations of similar intensity of meaning and value to the individual we might possibly expect to find similar reaction-patterns.

A partial explanation of the low intercorrelations may also be found in the differentiated system of values characteristic of any group of this age, which is related to and is sometimes the sole determiner of the interest in achievement in any particular field or specific task. Yet if this explanation holds, it is likely that in groups of individuals with less differentiated values, the more highly generalized interest in achievement will yield a more generalized aspiration level. In our experimental group we did find a number of boys who, responding more to the situation than to the task, desired to be "on top" in all the tasks; the classic remark to this effect was made by subject L64 who seriously stated, *"I would have worked just as hard if you asked me to push a peanut with my nose."* Others, spontaneously relating some of the tasks to fields in which they had experienced success or failure, in accordance with the nature of their reactions to these experiences, desired to do well only in one or a few selected tasks. Thus differing attitudes toward the tasks make what seem to be objectively one situation a number of different situations, in the psychological sense, both for different individuals and for the same individual at different times. We would suggest then that our data, rather than being a laboratory artifact, are in most cases a counterpart of the individual's differential reactions to different life materials which present a similar problem situation. Further, since overt behaviors may differ without an acompanying change in the basic motivation, average difference intercorrelations would then be the minimal expression of the individual's consistency of motivation in the experimental situation. An analysis of the interview data offers us some corroboration.

IV. ANALYSIS OF INTERVIEW MATERIAL

Even the most rigid attempt to restrict the preceding chapter to a discussion of the statistical results was destined to fail, since an analysis of the statistical data is inextricably intertwined with the material obtained from spontaneous remarks during the experiment and from the interview proper. The interview material appears to stand on its own, but one must recognize that its pertinence is in relation to the actual experimental situation, and it is likely that the extent to which this material is revealing of the inner motivation is due almost entirely to the fact that the experimental situation stirred up and brought nearer to the surface some basic attitudes of the individual.

As noted previously, the interview lasted an average time of two hours, and covered more ground than is indicated in this section. The following points are examined for the light they can throw on the function of the average difference score, the individual definitions of success and failure, and the relationship of the estimate to the deeper strivings of the individual: (*a*) Interpretation of experimental question; (*b*) meaning of the estimate to the individual; (*c*) chief factors influencing estimation; (*d*) explicit goal-strivings in this situation; (*e*) causes for feelings of failure; (*f*) reaction when performance score was below preceding estimate; (*g*) reaction to feelings of failure and (*h*) causes for feelings of success. These points will be discussed in a general way, and following there will be an attempt to evaluate the extent of the difference in attitude expressed by those differentiated quantitatively on the basis of the size of their median average difference score.

The following are some of the questions asked during the course of the interview, answers to which form the basis of this section.

> Did you always try to do better than you had done previously? Did you want to do better each time? Why did you (or didn't you) try to do better? Was there any level of performance that you reached that you were satisfied just to maintain—not go above?
>
> Did you at any time have feelings of failure? When? What was your reaction when you went below what you said you would do? What was your reaction each time you just made your estimate? What was your reaction when you went above your estimate? Did you at any time have feelings of success? When?

> Was the feeling of failure more intense than success, i.e., were you more depressed by failure than you were elated by success? In general, does the feeling of failure or success tend to spur you on more to do better?
>
> How did you interpret the question, "*What will you do next time?*" What did your estimate mean to you? On the whole did it represent the most you thought you could do the next time, or the least you thought you could do?
>
> What are you inclined to do under similar circumstances outside, e.g., if you meet some of your friends right before an exam and you're asked what grade you think you'll get?
>
> (If estimate was not also the goal.) Was your goal lowered simultaneously with a lowering of your estimates?
>
> What do you think were the chief things that influenced you in making your estimate at any time?
>
> What is your usual reaction to a feeling of failure?

The subjects were informed that there were no right or wrong answers to any of the questions asked so that nothing they could say would impress *E* favorably or unfavorably, and were asked consequently to be as honest as they could in their replies, since otherwise their answers would be meaningless, etc. As nearly as possible the questions were asked in the same way, but there were times when this was not possible. Not unexpectedly, the subjects differed in their readiness to answer some of the questions, as well as differing in degrees of resistance. When such resistance was met, the simple device of asking "why" after all statements that seemed to be merely evasive or stock replies, was successful in most of the cases. On the whole the subjects were delighted to talk about themselves at length, and proved most coöperative.

A. Subject's Interpretation of the Repeated Question Regarding Aspiration Level

We wished to understand the subject's general interpretation of our stock question, "*What will you do next time?*" On the surface, this appears to be a simple question of futurity which any "intelligent" person would react to with perfect objectivity. Actually the question of intelligence, within the average range at least, has no relation to the size of the difference score (5), and even when only futurity is implied to the subject, other factors immediately enter in to influence estimation in another direction. *H*45 provides a clear example. *"I took it literally—'How many problems will I do, etc.,' and of course there's a feeling that you don't want to say you're going to do worse and you have a feeling that you couldn't do worse. . . . I wouldn't let it come into my mind. . . ."* The interpretation becomes important, not as a clue to understanding the nature of the estimates, but for the light it can throw on the nature of the individual and thus indirectly on the difference score. In other than the "literal" interpretations one finds a tendency to interpret in terms that would put the individual and his performance in the best possible light. In some cases one notes that the interpretation is clearly a function of the individual's wishes and anxieties.

Thus, *"How did you interpret the experimental question?"* brought responses varying from *"I thought that you were trying to induce me to obtain a lower mark and I thought I could get a better mark than you wanted me to by my quota . . ."* (*M*20), throught the *"idea I thought I would improve and estimated I would improve with practice"* (*M*11), to *"Even if I thought I'd do worse I wouldn't say so . . ."* (*H*17). The answer given by *L*60 is extremely illuminating in view of the fact that one of his most outstanding characteristics is the attempt to avoid responsibility for the outcome of an event. He stated that the experimental question "could be taken two ways. The first one is the way I'm afraid it would be taken by you. Rating, seeing how a person rates himself and seeing whether he lives up to his rating. . . . But my way would be more startling if it could be proven. . . . I'd do it by rating, i.e., assuming no direct connection between what you can do and what you said. Sort of some people have insight to know, being psychic, and some people aren't psychic." The character of his rationalizations which are effective most of the time in preventing feelings of failure, partake of this same characteristic. This situation, which forced him to estimate a single irrevocable figure, was most irritating, and it is not at all surprising that we find him in the lowest median difference group. To estimate low is simply to run the least risk of being "shown-up" by poorer performance scores, and in addition there is a certain virtue attached to having performance scores higher than estimates. That the interpretation in itself had little influence on the nature of the relationship between performance and estimate, is clear also from the fact that many subjects admitted to reacting first, and only thought about an interpretation when asked. At the other extreme we have a subject (*M*58) who interpreted the question in one way and deliberately did not follow his interpretation because he was "over-cautious and tried to evade anything." *L*78 expressed somewhat the same attitude in answer to "How did you interpret the question, '*What will you do next time?*'" He replied, "I'm quite positive you were trying to correlate my different answers and I guess I wanted to be sure I didn't make it too high or too low—sort of strike a mean—I mean possibly I wasn't being candid in my answers. . . ."

In view of (*a*) the varying interpretations found within the group, (*b*) the occasional disregard of one's own interpretations, (*c*) the presence of interpretations which were clearly related to the individual's anxieties and insecurities, and (*d*) the fact that interpretations often arose after the reaction, we are not inclined to find any significance in the fact that more persons in the *H* difference group than in the *M* or *L* difference score groups regarded the estimate as a goal. Those who did so in the *H* difference group may simply have been rationalizing the disparity that actually existed between their performance and estimates. One of the subjects in this group, (*H*19) was sufficiently aware of his motives to express his reaction in these terms. He said, "The mark I tried to get was pretty honestly what I thought I would get but I probably rationalized going below [it] by saying estimate was a goal. But I hadn't set out considering it a goal." ' *H*76 interpreted the experimental question in terms of "perseverance" instead of "goal." There is reason to believe that this was sheer wish-fulfillment, a rationalization of his persistence in maintaining rather high estimates, which serves to place his behavior in a favorable and "commendable" light. For these reasons then, the clue to the differences which made for quantitative differentiation of the three groups, is probably not to be found in their different interpretations of the question.*

B. MEANING OF THE ESTIMATE

When the meaning of the estimate to the individual is analyzed, we see clearly that not only does the explicit estimate not present

---

*Current work with about 30 subjects indicates no difference whatever between response to "*what will you do next time?*" and "*what do you intend to do next time?*"

an accurate picture, in many cases, of the momentary level of strivings, or aspiration level, but that the *Aufgabe* is perforce not the same for all individuals despite the fact that the general laboratory situation and experimental question appear to be controlled variables. By restricting the definition of aspiration level to the estimate which the individual "*explicitly* undertakes to reach" (1, p. 119), the *implicit* level of strivings, which in the last analysis determines feelings of success and failure, is either overlooked by Frank or assumed to be given direct overt expression.

The different meanings which the estimates had for any person could be roughly classified under (*a*) the *average*, i.e., the actual score expected the next time, or the mean between the least and most expected, (*b*) the *least*, i.e., the lowest performance expected, or even a score below expectations, (*c*) the *most*, i.e., the highest score expected, or even a score slightly above expectations, and (*d*) the *mixed* response, i.e., where the individual felt he had utilized the estimate differently at different times. That these classifications are only superimposed on material differing qualitatively, can be seen from a sample of the responses given below.

(*a*) The *average*.

M73 . . . when I said 34 that's exactly what I thought I would do, neither the least nor the most.

M40 I think they were just about average. . . . I usually allowed for a certain increase *but I felt I could do more* than that. . . .

H38 I tried to determine what I could do after the first few trials and then let that represent an average. . . . Goal was always a few points above the estimate.

M50 It was just about what I thought I could do next time. I hate to go too high, but still I like to go high. I hate the idea of feeling, "Gee I'm not improving." . . . I suppose my real goal was beyond my estimate. I felt, "Well I could reach this next time and next time go beyond it." I didn't want to overestimate too much.

(*b*) The *least*.

H45 I'd say minimum rather than maximum. I never made what I thought was a foolhardy guess. I took things into consideration, what I thought was the normal rate of proficiency. . . . Of course, there's a feeling that you don't want to say you're going to do worse and you've a feeling that you couldn't do worse. . . .

L52 Estimate was the minimum I would have liked to achieve.

L9 I got no personal satisfaction out of getting 35 when I said 30 because I always felt I could do better.

L54 I think I tried to make it as the least, sort of, "Well if I did better than that I'd be pleased and surprised."

M69 Usually estimate was the least. . . . Because if I think I overestimate I probably would be disgusted when I finished by missing it by a wide margin—just a general disgust with the whole test idea and I wouldn't be able to do as well as I could. . . . It seems sort of silly to estimate more than you think you can do. . . . My estimates were the least because though I knew I might fall below I was quite sure I'd make it and possibly pass it. . . .

M75 I always had a feeling I could soar to great heights above my estimate if I had an easy set of questions.

(*c*) The *most.*

M57 . . . when I set the goal [estimate] it had been slightly higher than I expected, but I hoped that the added incentive would help me reach it.

M37 Most of the time it was the highest I thought I could do and sometimes it was that plus a few more to aim at. . . . When I lowered the estimate it was possibly what I thought I could do but it wasn't necessarily a lowering of my goal.

L77 Thinks estimate represented the most he expected to do, although, "I should say that most of the estimates were arbitrary." (See analysis of case below).

H33 After first two tests estimate became a mark to aim at. "I wasn't sure I could do it but I felt I could come nearer to it than if I gave a lower estimate."

H17 Probably the most I expect to. . . . *Even if I thought I'd do worse I wouldn't say so. I don't see how you could do worse anyway, unless your pencil slipped or something.*

(*d*) *Mixed.*

M61 It varied with the tests. In one test it may have been that I hoped to do better, in others I thought I would do better, and in others it was a wild guess based on the previous score. (n.b. Latter restricted to the synonym test). . . . In steadiness I thought it would be the least improvement I could show. In the target it was the least. In the synonym tests I would say

> it was the most. In math it was the least. Most of them were the least. In one or two cases on the target it was what I *hoped I would make, rather that what I thought I would* make. Today [s-d, add, and cancellation] I judged on what I thought I could do. . . .
>
> *M69* Most of the time the estimate meant that I intended to try harder the next time and go forward and sometimes I didn't care about it and said the same number because I felt I was concentrating too hard and wouldn't do so well. (n.b. Strange explanation for something one does not "care about.") Quite a few times I didn't have any definite feeling about it. Just said a number I thought plausible. . . . If I thought I could do better I made a higher estimate. If I thought I had reached the height in that particular thing I just didn't bother.
>
> *H72* I tried to adhere to the lowest possible minimum I thought I could achieve. Maybe in some cases it was the maximum. . . . Something I was sure I could achieve, in other words, without too much effort. (n.b. This subject has the highest median average difference score.)
>
> *H15* I set it as a goal I might strive for, except in certain cases where I had a feeling of failure, I put a score lower than . . . I felt I might do.
>
> *L16* Sometimes I purposely put estimates low so as to get them but not consistently. Also—for the most part they were what I thought I would really get. I tried to make it a habit not to expect too much so I wouldn't be disappointed. . . . That's a philosophy in my home.
>
> *L60* It alternately meant what I thought at the minimum I could do and next time it might have meant what I hoped I would do and another time just what something told me I would do. About 50 per cent of the time it was the minimum I thought I could do and 25 per cent for the others. . . . Also sometimes I set the estimate ridicuously low so I was sure I would go over it. (n.b. In the s-d and cancellation tests.)

At least until we discover what exact relationship, if any, exists between an individual's implicit motivating level of goal-striving and his explicit level, it seems necessary to concentrate attention on the former, since the latter serves quite different functions for different individuals. As seen, the estimate might represent a level of performance the individual undertakes to *overreach,* while to others it might represent a mark they hope to come close to (actually being prepared not to attain it), and to some the meaning of the estimate may change in response to specific experiences. It is clear then, that quantitative differences in the size of the difference scores can give us no accurate picture of possible differences in aspiration level, since the actual level of momentary strivings may be of the same magnitude in individuals widely divergent in difference scores. There is some indication, however, that more individuals in the *L* group regarded the estimate as representing "minimum" achievement than in either of the other two groups. This, as will be pointed out later, is related to the fact that in the *H* group the estimate tended to function as an incentive whereas in the *L* group it was more intimately associated with feelings of success and failure.

Explicit estimates yield more rapidly to changes than does the motivating level of goal strivings, although in certain cases the explicit level is also that which is striven for. If an implicit level of momentary strivings is not revealed via the explicit aspiration level, it is easy to understand why in many cases, the implicit level is more resistant to change than the explicit estimate. The explicit estimates are forced to keep in closer contact with the real performance scores, since the estimate, regardless of its meaning, nevertheless represents a *commitment* made to another person, and whatever one's attitude towards the estimate, one feels compelled at least to come close to the actual performance. As one subject expressed it (*L*14), "No matter what you say within reason I think you expect to do it whether you want to or not. Otherwise you look a little foolish if you keep setting figures and then don't make them." This explains why explicit estimates, even though starting out coincident with implicit inner strivings, may become distinct in the course of a single test. E.g., when performance scores drop for one or a few trials, it is likely that most estimates will decrease accordingly, while the original momentary level will probably be maintained as an inner striving which the individual attempts to reach, regardless of his estimates, and regardless of whether he is aware of the inner level. On the other hand, if performance scores should suddenly increase, in some cases where the individual is not *convinced* that the increase will continue, his explicit level may remain the same while the implicit level is changed to a higher score.

This experiment offers no evidence one way or the other as to the existence of an aspiration level *per se*, but the evidence obtained here does point to the fact that the estimates given in this and similar situations do *not* accurately express what the individual is *striving for.* It is possible that from these data one could get a crude idea of the differences in actual aspiration level by analyzing the individual statements for an indication of whether the actual level of momentary strivings is above, coincident with, or lower than the explicit estimates. Unfortunately too few subjects made sufficiently definite statements to warrant such an analysis. In some cases, moreover, such a surface analysis would be misleading, since there is reason to believe that even the conscious level of strivings and expectations might not accurately express the level of performance towards which the individual is actually working.

Many of us are sometimes fleetingly aware of an inner ambivalence towards the outcome of some activity engaged in, i.e., we want to, will try to, and somewhat expect to make out well, and at the same time we "don't care" and don't expect to make out very well. Our surface behavior may be a reaction of bravado towards the outcome, or it may take the form of low explicit estimation, or we might poke fun at the task, mocking and belittling it to the effect that any outcome would be considered of no consequence. Yet the actual level of striving is high. One must recognize, of course, that awareness of ambivalence is not necessary as evidence of its presence. Similar behavior may occur when awareness of the ambivalence and the high level of strivings are repressed, and only the low expectations come to consciousness. In some cases, the conflict occasioned by ambivalence may be solved by the repression of the low level of expectation and striving. Evidence for the functioning of the latter may perhaps be more difficult to obtain than in the case of repressed high level of strivings. Perhaps the main source of evidence in such cases would be the expression of anxiety, "incompatible" with the conscious level of strivings, towards the outcome, and sudden disproportionate decreases in estimates following an experience of failure. Some in the *H* group behaved in this fashion, and it is perhaps in this light that the following statement of *H*28 may be most clearly understood: "Consciously the estimate is also what I expect but unconsciously I don't expect to get that high." Of course if that feeling were "unconscious" he would be unaware of it, but the statement probably is descriptive of a glimmering awareness of his ambivalence towards his future achievement, where the "low" expectations were somewhat inhibited. *H45*, referred to previously, who would not allow the thought that he might do worse come into his mind, provides another illustration.

There are several cases in the *L* difference score group, where it appears to the writer that one finds generalized personality systems organized around the repression of high levels of striving, or rather that this repression is symptomatic of a certain kind of personality organization. I.e., there seems to be a generalized attempt to seek a level of reacting which would provide, so to speak, least grist to the mill of insecurity, and allow the individual to function under the least possible strain. A full discussion of this statement would lead us too far afield into depth-psychology, but a brief presentation of one case in the *L* group will illustrate what is meant.

L53 does not feel "sure" of himself "when it's something new, but on the whole I do." This probably means that he would tend to avoid wherever possible the "new" and unexpected, since the anxiety feelings would be most acute under such conditions. A conscious level of expectations, restricted to the minimum he might possibly expect to attain, would best fit in with such an orientation, since such a level would be least conducive to arousing insecurity feelings regarding future achievement. We expect him to feel then that "When I fall below what I aimed at it bothers me," since failure is experienced thereby and feelings of failure feed the fires of insecurity and anxiety. This low conscious level of expectations provides a bulwark, but if it is unduly low in reference to the actual achievement, its value is considerably diminished. Thus, "I was a little bit surer of myself when I didn't go up so fast. . . . I wasn't *sure* of how much I would improve and therefore I would anticipate only a little higher when I did give a higher estimate; but if it went back I would more likely estimate lower and start all over again." We can readily see that his conscious level of expectations may not be actually revealing his level of strivings, nor can we be sure that his conscious level of expectations is the one that is really motivating his behavior. We have no clue as to what level of performance he is actually satisfied with, though it is always possible that any level above his minimum conscious expectations may be satisfactory. There is certainly no way of judging this from his explicit estimates, since he as well as others in the *L* group stated that they always strove to increase their performance scores, although such strivings frequently were not mirrored in the estimates.

The subject's tendency to make and stick to first impressions may be seen to fit very well into a system of avoidance of anything leading to uncertainty and insecurity, and leading away from "safe" predictions. This reluctance to change first impressions demonstrates the widespread nature of his attempt to stabilize and "put in a safe place" every possible feature of his environment. Thus it is that he finds himself resisting facts, when subsequent information about his acquaintances forces him to change his opinions, the direction of change, from "like" to "dislike" or vice versa being quite irrelevant. Although he attempts to rationalize this attitude and points out that after all what originally gave him that impression should be sufficient evidence against changing, his final statement is the crucial explanation: "I just don't like to change." His behavior in this laboratory situation is clearly related to this general characteristic.

Since a score below his estimate would be productive of too great anxiety, the response necessitating least change, would be the maintenance of low estimates, and as further protection, low conscious expectations. In extra-curricular activities, we learn without surprise that he would rather see "someone else lead an organization and follow along. . . . When somebody else does, I feel I could have done, not better, but just as well." He admits that this attitude is partially due to a fear of failing if he did go in and do it wrong, but more important is the "*uncertainty* as to the outcome." Finally we have evidence that the systematic seeking of a level of being that would both tend to allay existing tensions and be least likely of making demands on the established psychic equilibrium, has been developed into a philosophy which seems to rest upon a "categorical imperative." Towards the end of the interview which was devoted to a discussion of the basis of his estimates, he attempted to justify his behavior. "I don't feel anyone should overestimate themselves." . . . *E: Why would considering yourself "good" be overestimation? Your scholastic grades are above average.* "Probably because I don't consider myself good. . . . That's the way I look at life. I don't consider it right [to consider oneself good]. . . . That's the way I think it should be. . . ." It appears then, that the conscious level of expectation and striving, which in this case is coincident with the explicit estimates does not represent his actual level of strivings. We believe that the fear of not being able to achieve the actual level of strivings created a conflict which was solved in this case by repression of the real momentary goals and creation of a conscious level of strivings which he could be sure of always achieving. His behavior must be understood, we believe, in the light of the repressed rather than the conscious strivings.

It appears not only that a lowering of conscious expectations is sometimes brought about as a result of the individual's inability to face potential failure, discouragement and so on, but also that the same condition may be brought about as a result of finding that this conscious restriction of expectations interferes least with proficiency. E.g., if an individual is inclined to become very much upset when a given performance does not come up to certain standards, his succeeding performance is likely to suffer in quality, and after a while the standard may be lowered to offset any possibility of going below it in achievement. If an individual's feeling of insecurity becomes aggravated by uncertainty as to the outcome of any event, it is likely that "high" estimates will be avoided because one can be less sure of doing one's best than of "not doing less than" a particular score. This anxiety reaction to "high" estimates differs in extent and intensity in different individuals of course, and is not restricted to those persons who usually estimate "low." Such anxiety feelings have been revealed by individuals in all three quantitatively differentiated groups, occurring most frequently after the estimate has been raised considerably above the preceding one, which apparently arouses such tensions as to interfere with the performance and lower efficiency to a point below that of preceding scores. Some of the subjects to whom this happened, became aware of the effects of tension, of "pressing too hard," and reported that at such times when the performance score dropped unduly, they deliberately decreased the estimate to one which they were sure of at least achieving and probably surpassing, because they felt that if they "relaxed" they would get a higher score.

It may be argued that since the effect of "pressing too hard" was to decrease the performance score, and that since a decrease often brought about a feeling of failure, the decrease in the estimates was simply a response to the lowered performance score, and the reasons offered for the lowering of the estimate merely rationalizations. If the effect of decreasing estimates was fortuitous, and the reasons simply rationalization following realization of the effect, this would not detract from the validity of our assumption. The first lowering of the estimates in response to a feeling of failure (following a sudden decrease in performance due to anxiety feelings), may have been accidental in the sense that the individual was not accomplish*ing what he set out to do by the methods he was using; he changed* his method, i.e., decreased instead of increased his estimate. After doing this, the desired performance level was more closely achieved; the individual became aware of the advantages of such a change of tactic under those circumstances and deliberately used it under succeeding similar conditions. When the question was brought up during the interview, the subject rather naturally described the situation as if it had been deliberately brought about. Corroboration for this view is obtained from some subjects who withdrew from the tension-situation differently. *M*82 reports that during one trial on the addition test he became very anxious about his score, worried lest it not compare favorably with others, afraid that it would be even lower than his preceding score, and since these anxiety feelings were probably intolerable, he withdrew from the field entirely, deciding that the test wasn't important after all and it was silly to worry about it. Having thus relaxed, he was able to work faster and achieved the highest score by far that he had yet done. The connection between his relaxing and getting high scores struck him as singularly a cause and effect relationship, and he tried not to "care" for the remaining trials, with fair success. Another subject whose performance after the first few trials on the target became consistently worse, tried first to prove that the discs were of uneven smoothness and that failing, on the 8th trial suddenly announced, "You know I'm not interested in this," whereupon his tenseness disappeared and his performance score, like magic, rose to an unprecedented height for him. That the performance scores did not decrease with such an emergency decrease in estimates or an explosive "loss of interest" seems proof that the actual level of strivings and effort to reach that level, whether conscious or not, did not really alter in these cases.

Such an attitude towards future achievements, brought about only in critical situations, may in certain cases develop into a systematic approach to all untried and unpredictable undertakings. The genesis of this approach may not be known to the person, yet it may be maintained because it is at that level of reaction that he feels most free of pressing fears and tensions. The reasons offered by some for the maintenance of both "low" estimates and "low" conscious levels of expectations and strivings support this view. In *L77* we have reason to believe that, although he regarded his estimates as representing the "most" he thought he would achieve the next time, his conscious level of goal-strivings was restricted to the "least" he might reasonably expect to do. Instead of estimating low in life-situations, however, he assiduously avoids making any estimates at all.

> I feel if I set myself too high a mark I'll be disappointed when I don't get it and if you set too low a mark I felt I might be discouraged. If I say I won't get over a *C* I might be discouraged and say, *"What's the use of studying if I'm going to get a rotten mark"*—so I just tell the boys who ask me before an exam to wait till it's over and I just study and don't worry about how well or how poorly I'm going to do.

We learn, moreover, that he is "satisfied" with a *B* average, and in view of this discussion it is not surprising to discover that practically no mark is ever below a *B*, that in fact his scholastic average is *A—*, which he hastens to assure *E* has nothing to do with his being satisfied with *B*. That this conscious level of strivings or aspirations is in the nature of a protection against discouragement, against potential failure, seems clear from the following statements. "I'm not going to set myself an *A* standard and slave away at it and if I don't get it get completely downhearted about it." In answer to the question of whether he wanted to be an *A* student or not, he replied, "I would say that I want to be at least a *B* student, how much above doesn't matter. . . . Of course the better I am the more I like it but if at the end of the year I came out with all *B*'s I wouldn't break down and cry about it." In this experimental situation he also avoided setting any "goals" that were meaningful to him "except to do better than I did the last time." His explicit estimates he characterized as being rather "arbitrary," often not even indicating that he was striving to do better. The reasons offered for this behavior are similar in nature to those quoted above.

> I tended to base my estimates practically entirely on previous performance. . . . I didn't think about doing anything else. I

> didn't like to set my estimate too high because I felt it might make me do worse than if I had set it a little lower—possibly by trying too hard.

Finally, the fact that he only experienced success when his performance was *quite a bit above* the estimate, seems conclusive proof that his conscious level of strivings and expectations is restricted to "minimum" expectations, and that he is actually motivated by a higher repressed level of strivings. At the end of the interview when asked whether he was inclined to "play safe" at all times, his reply came quickly without hesitation, "Absolutely—that's probably the most characteristic thing about me, slow and sure. I'd never make a gambler."

The frequent dissatisfaction with obtained performance scores by individuals differing in the size of the average difference score might also be an indication of either repressed momentary strivings that are higher than the conscious level of strivings, or just the lack of correspondence between level of estimates and implicit level of aspiration. If this is true, then in those cases where there is reason to suspect repression of certain motives, the repressed motives appear to be more influential in determining the individual's subsequent behavior and reactions than the conscious level of momentary strivings. Degrees of repression, or the clarity of awareness of one's motives probably bears little relation to the strength of the repressed or otherwise unexplicit motives. Horney confirms this point. "The degree of awareness varies. A person may be aware that anxiety prevents him from performing tasks in a satisfactory way or he may only have the feeling that he is unable to do anything well" (8, p. 57). An individual may consciously set estimates lower than he expects and be at the same time aware that he is striving to reach a higher unverbalized goal, or he may not be aware that he is striving for any particular score and only conscious of dissatisfaction with all scores not above his explicit estimates. In those cases where the unverbalized level, conscious or no, is lower than the explicit estimates, then feelings of success are often experienced when the performance is below, but almost at the level of the estimate. In this light we also understand why *L*79, who insisted he really expected to do poorly, agreed that if his life depended on the accuracy of his prediction, he would raise his explicit level of estimation.

This discussion, attempting to point out the intricacies of motivation involved in the whole question of level of aspiration, and presenting some case material to demonstrate that even a knowledge of the inner conscious level of momentary strivings, let alone the explicit estimates, may not reveal the level of goal-desires which actually motivates the individual and is the chief determinant of subsequent behaviors, is not to be construed as implying that people do not differ in heights of aspiration level. Nothing that has been said is incompatible with the probable facts that people do differ—even with respect to their levels of aspiration. Nor has it been implied that all "low" conscious levels of momentary strivings are indicative of a repression of "high" goals. Rather, we have tried to show the hazards involved in tapping deep-lying structures with surface-scratching scalpels, such as the present aspiration level technique provides. This discussion has, of course, no relevance to those persons who were not genuinely *involved* in the task. It should be pointed out, however, that a casual treatment of a given material is not always indicative of a lack of concern with it, that "dislike" of a thing is recognized as often being "the means of avoiding and the consequence of the anxiety" (8, p. 58), that there are more ways of avoiding failure than repression of the fear of failure, and finally that there are ways of expressing conflict other than by ambivalence.

C. CHIEF INFLUENCES DETERMINING NATURE OF ESTIMATE

It has already been shown that estimates though influenced by the reality principle, i.e., previous performances, are nevertheless, within differing extents for different individuals, subject to the inner demands of the individual, such as anxiety and insecurity feelings, desires to excel, to succeed, to avoid failure; and to the actual level of momentary strivings, the disparity between this level and the explicit estimates, and general past experiences which have helped determine his particular personality organization and thus his reaction to such demands of both inner and outer forces. The following quotations have been chosen to illustrate the influences of these factors and to show the variety of such influences. It is clear that the different behaviors of the same individual are not really "inconsistent" with his motives.

> *L70* "I rather doubted I would get up higher [than estimated] and once I said something I stuck to it. . . . Also afraid

I couldn't make exactly what I'd say. . . . Had thought I might go downhill at any moment from taking my attention off. . . ." "High" scores are not responded to as evidence of ability but as a good event which he had had nothing to do with in bringing about. Did not experience any feelings of success in this situation, though he did get feelings of failure when his performance was "lower than I thought I could and should do." Has very little desire to "excel" in this or life-situations. In general has few feelings of failure or success. Has a B+ average and claims he is satisfied with it, and though he is trying to do better than that he is quite sure he cannot because he has not done so up to now. Did not seem to be involved in this situation other than trying not to get scores below average.

L79 Influenced by previous performances and "whether I felt I was going to be better or worse." Half the time felt he would do worse and the other half felt he would do better the next time. In general suffers from a lack of self-confidence. "I was always afraid I was going to slip or something, lose a lot of time and fall behind the previous score." Admits, however, that if his life depended on the accuracy of his prediction, he would estimate higher, since usually he does better than he estimates and not infrequently has ambivalent expectations toward his future achievement.

L52 Since did not know what the average score was, in some cases did not know whether to be encouraged or discouraged by his scores. When he was temporarily discouraged by a decrease in performance or a score below his estimate, "there was a thought of scoring higher and coming back and some of them when there was a low score there was a loss of interest." When he lost interest, effect on estimates was "I set a number I thought I'd be able to beat in a way to give myself encouragement." This he did in three of the tasks when "the tests went down a good deal, (I) set a number which I was pretty sure to beat next time—to show I was starting upwards again." There is a notable exception to this behavior, namely, in the addition, which he claimed was the most interesting to him. There, "I was after a certain number, whether I made it or not I was going to try for that in the next shot." Also claims that he always tried to give an estimate that he was sure he could at least make and probably surpass. That his estimates were not always the minimum was clear to him, however, and he reacted by almost instantaneous rationalization to every occurrence of a performance score below his estimate: "Well that's what I would have *liked* to have made, not what I *thought* I would do." He recognized that actually he had *thought* he would at least achieve the estimated score.

L22 "Influenced by my feeling whether I felt I was getting along well and whether I was feeling tired." Was very cautious, always decreased estimate at first sign of a decrease in performance because afraid it might continue to go down. He decreased estimate hastily to the level of one of the lowest obtained performance scores, even thought that score might reasonably be considered a "fluke." He did not want to be accused of overestimation, in fact would have preferred to be accused of underestimation. Claims that he did not at all times think he could beat the previous score. Also sometimes felt he was concentrating too hard and if he'd just relax he could go faster. Since he did not know average performance score, took his cue for judging his scores from the first performance score in each task.

L47 "I don't think I ever raised my estimate very much. I gave myself a little margin of safety. . . . I always put down what I thought I'd get next time. Sometimes maybe I put in what I hoped I'd get—that may have crept in some place. . . . Estimate probably the minimum. I tried to pick out what I would make but it was probably the minimum. . . . If I did fairly well in one, say 35 and aimed for 40 and got 30 and said 34 and if I got 34 would I be satisfied with 34? Yes, I'd be satisfied with 34 but in answer to the next question I'd say 40. . . . Funny if I'd made 45 the next time I'd only estimate 40. . . . Just naturally cautious I guess. I mean if ever I'd get out on a limb or a little higher than I ought to be." *E: You mean you'd always like your performance to be a little higher than your estimate?* "I guess that might be it—but not consciously. Consciously I thought, 'I'm a little bit above my head and I better not estimate too high because it probably won't happen again. . . .' "

L78 ". . . As I explained before, I wanted to avoid being too far above or below my actual score. I tried to come as close as possible. . . . *I was probably trying to avoid having you think me either abnormal or subnormal.* I just wanted to be average. . . . In short I probably think I have more than average ability and I wanted to cover up the fact that I thought so from you. . . . When I'm not sure, when I don't know whether I'm going to make a high or a low score, I tend to strike a mean. When I'm fairly sure of the test than I can allow myself to estimate lower [than he expects or is striving to do]. [Also influenced by] previous performance and my general condition, i.e., whether I feel my hand is getting tired as in the steadiness." . . . *E: Why did you feel freer in naming higher estimates in the target?* "Probably because it has nothing to do with intelligence. It has less of any inward feeling, the answer seems to be more spontaneous, i.e., it seems to be less important. That seemed to me the only test I could go at with a marked degree of confidence. . . . These other tests I never felt quite sure of myself—that I was giving an adequate performance." In the addition test he had a very low minus average difference score and was questioned as to the reason for it. His reply is quite significant and not unexpected in the light of the insecurity implied in the above statements. ". . . I'm making a statement of fact that the test I felt most secure in is addition." It is also significant that he did not get a feeling of success when his performance was above the estimate—*"Possibly because I had a feeling I gave a low estimate simply so that I could do better so that it didn't give me a feeling of success at all."*

M20 ". . . See if I had said 40 and my average was around 35 I wouldn't get 35 but 33. I'd fall below, but if I had said 35 I would have gotten 35 or maybe even 40." *E: You mean that setting a high goal so unnerved you in trying to make it that you would tend to make out more poorly?—But if you set a goal which you could be sure you would at least reach, if not do better than, then you would be able to do more?* "That's it exactly—with but two exceptions, the cancellation and steadiness which I felt I could reach whatever goal I set out to."

M61 ". . . I felt rather than set a high standard and fall below, I'd rather set a low standard and try to go above it." Also when he did not do as well as expected or estimated he "lowered the standard a little and tried to go *above* that standard."

M62 "I usually set it (estimate) so I had something to work towards. . . . I did not know what others had done but I was trying to rate highly. . . . Yes I always thought I could do better. I never knew when I had reached a peak or anything. . . . Well in the beginning I really thought I could do better, but after a while towards the end I guess I rated myself higher because I wanted to do better. . . . My estimates were never too much above or below, so I didn't have to lower them much [when performance below estimate]."

M74 Attitude towards basketball estimates in life acted to make for relatively high estimates in target. Accounts for estimates higher than performance in cancellation and addition as follows: "I never felt any intense concentration. I just felt I could do better." In synonyms and s-d, the "learning" test, estimates were very close to actual performance, which may be accounted for by his statement to the effect that where his personal pride was involved he usually set estimate at a point he felt he would surpass.

M50 "Sometimes I felt like putting the estimate very low, so as I could at least guess my ability. . . . In these tasks I was pretty sure I could do better and then I sort of reached the place where I'd rather have a little higher goal to shoot at, than put it down low. . . . I'd rather try for something up high and miss it than try for something low and just get it and sit there. I don't want to just try for a C or a B—. I'd rather try for an A and miss it, get a B and I'll be satisfied with it. . . . I figured out what I had done and figured I'd put a little above and if I did get higher, I put it up above and figured I would try harder. Some things I realized you couldn't improve much and if I improved one point, I felt that was pretty good."

M32 "I always tried to exceed my prediction—I never remained satisfied. You see I set a high predictive level and tried to reach it because I found out from life that if I do that I get good results." If these two statements are not contradictory then they would indicate that his explicit estimates were lower than his inner "high predictive level."

H18 "Wanted to keep my goal as high as possible but not always as high as I would have liked to have gone or thought I could reach."

H29 Wanted not "to appear dumb."

H45 "I don't really think I ever had a feeling I was going to do worse. I wouldn't let it come into my mind and as a matter of fact I don't think I ever said 'the same' because I thought with added experience I should become more apt. . . . And secondly a desire to show you a feeling of self-confidence."

H23 "We're never satisfied with what we have so we keep trying to do better and better. . . . Have to keep goal higher."

H72 Desired to aim high, to excel, to succeed.

> *H65* "If I didn't do as well as I did before I was inclined to go wild. If I hit 30 and estimated 35 and hit 28 I would probably go back to 35 or even a little higher. It was just that I had failed and was annoyed with myself and determined to overcome losses. But if I did better than previous performance I was inclined to go higher but not to any wild extent. . . . I just thought, 'Now I'm getting used to it and I'll make a higher estimate.' Also knew it would be impossible *to make too big a jump. . . .* At times I did *take* a chance, but frequently I was quite conservative, quite variable I imagine. I wanted to appear conservative—not too wild, although it didn't hold true for all the answers I gave. . . ."

In discussing the average difference score and the problem of "generality," it was suggested that low difference score intercorrelations did not prove the absence of generality of motivation, or even of behavior varying in response to the particular motivation. It was pointed out that any overt behavior did not indisputably indicate the presence of a particular motive, but rather that any given overt behavior might be manifested in response to quite different inner needs. It was felt that the corrollary of this was also valid, i.e., that quite different overt reactions might be elicited in response to similar needs. In regard to the problem of "generality" this meant that congruent and generalized attitudes toward particular aspects of the environment might be displayed, although through the arousal of different attitudes toward different events, the crude quantitative measures would appear to indicate "specificity." This approach to the problem of "generality" which recognizes the significance of "polarities" and finds it not inconceivable that the same factor can be responsible for two quite different behavior-resultants, should find corroboration or denial in the survey of factors influencing the nature of the estimates.

At first glance it would appear that individuals in the *L* group evinced, on the whole, much greater insecurity in regard to their future performance and more often anticipated that their score would decrease rather than increase in the following trial. Sudden increases in score are more often regarded as "flukes" by this group and as genuine increases by the *H* group, and sudden decreases in performance are regarded as accidental and easily explainable by more in the *H* group than in the *L* group. The very difference in the size of the average difference scores in the two groups would indicate this. Can we say, then, that there is a one-to-one relationship between the size of difference score and amount of self-confidence?

There is no doubt that the factor of confidence influences the nature of the difference score—but not in any simple one-to-one fashion. In the *H* group we find several who wanted to impress *E* with the appearance of self-confidence, two admitting that though they did not always feel confident they would never admit that they felt they might do worse the next time. In fact, the thought of doing worse would not even be allowed to enter their minds, thus proving at once both the presence and immediate repression or rationalization of self-doubts. *H*29, on the other hand, gives evidence of being more aware of his insecurity, the "high" estimates being in the nature of bravado. *M*20 revealed his insecurity, not through "low" difference scores but through his intense suspiciousness of *E*, which took the form of projecting on to *E* his fear of not doing well, i.e., he felt *E* was trying to make him do poorly. It was this consuming insecurity which he attempted to hide both from himself and *E*, and his desire to do well: *"I didn't want to prove myself dumb"* (which in its very wording is suggestive of his attitude), that drove him at one time to "overestimation" and another time to slight "underestimation." The following picture of *H*23 is hardly suggestive of self-confidence:

> I think that falling behind spurs me on more—gives me a feeling of insecurity and I made a greater attempt to obtain a better score. . . . I'd like to [be an *A* student] of course. But I have a bad idea that I've bitten off more than I can chew. Sometimes I think I'm worthy and sometimes I think I'm unworthy. . . . I know which [feeling] is more persistent—that I'm not worthy of being here. . . . I guess you couldn't be too ignorant and get a B+ in Columbia College—so I guess it's all my fault. Anyway that's the way it looks. I guess I shouldn't feel that way. . . . You see I go up and down and if I go down I say I'm not any good and if I do well I say I'm pretty good.

He has the second highest median difference score and is third lowest in variability. Only once did he lower an estimate and that was "because I felt I was pressing too hard and if I did lower it I would make that score or better." *H*33 admitted doubts concerning his ability but he was terribly anxious to do well and impress *E*, and though he was never sure he could achieve it he set estimates high because he felt he would have a better chance of approaching them than if he gave a lower estimate. At the other extreme we find *L*64 who displayed singular self-possession in overt behavior. He claimed,

> I've pulled myself apart pretty thoroughly. I think I'm pretty good. I know I don't do certain things well but I don't feel particularly badly about it. I'm not particularly outstanding in anything except my music and there people tell me I am outstanding. I don't know whether I am—but there's nothing to be blue about. . . .

He estimated low because he felt that "when one is estimating one's performance one should be as mathematical as possible. . . . In most of the tests I had no idea what you were driving at or what they meant so all I could do was try to estimate as closely as I could." *L*78 explained his estimates in terms of self-confidence.

> In short I probably think I have more than average ability and I wanted to cover up that fact from you. . . . When I'm not sure, when I don't know whether I'm going to make a high or low score I tend to strike a mean. When I'm fairly sure of the test then I can allow myself to estimate lower (than expects to do). . . . *Yet I always think to myself that I'll do better than the answer I give.* . . .

We could quote other instances, but these cases provide sufficient illustration of the fact that though the lack of self-confidence may be responsible in some cases for the feeling that the next performance will be worse, and for the fear of estimating "high," in some cases it is responsible for "high" estimates, since the latter may serve to bolster the individual's confidence or be used to hide from *E* the subject's awareness of his lack of confidence. Similarly, a feeling of confidence in one's ability may act to maintain estimates higher than preceding performance scores since the individual never doubts he will do better the next time; or it may take a more perverse form of estimating lower than expectations and strivings, on the basis perhaps, that a firm belief that performances will be higher than the estimates serves to increase confidence rather than (as in some cases) be a reminder of one's doubts. Finally, in some cases, the factor of confidence may not influence the nature of the estimates at all. Other factors such as interest, meaning, goal-strivings, reaction to feelings of failure and success, may be more primary.

These quotations also indicate that differential interest in the tasks themselves may make for a difference in the character of the estimates, quite apart from the fact that lack of such interest was frequently accompanied by a genuine striving to do well simply for the sake of coming out "high," or "to save my face," or in order not "to prove I was stupid." The interview material brings to light the fact that interest in a task makes for a wider divergence in the kind of responses elicited by the aspiration level technique than tasks lacking in special appeal. No single factor, however, can be held accountable for any obtained results in a group of individuals, although in certain cases it is possible to find one particular characteristic or attitude chiefly responsible for the resultant difference scores. The same attitude on the other hand, can be found in individuals differing markedly with respect to the extent of their average difference scores.

D. Goal-strivings

The material for this question was obtained most indirectly so that a comparison of the three groups on the basis of the answers cannot be considered conclusive. The direct questioning was somewhat as follows: *"Did you feel you had to do better each time?" "Did you want to do better?" "Why?"* If the subject began by saying "It's natural," or "instinct," or gave other such non-personal responses *E* kept asking *"why"* until she felt that the subject had revealed as much about his motives as he was able to do. Thus all references to wanting to come out "on top," or to "beat" the next fellow are in a sense quite spontaneous. A bird's-eye quantitative analysis of the answers, samples of which are given below, does not reveal any basis for differentiating the three groups.

(*a*) *Wanted to come out on or near the top.*

> *H76* Thought maybe I wasn't doing as well as others and therefore wanted to do better. . . . Rising to the top [in life] is a thing I'm continually thinking of.
>
> *H24* . . . Just the fact that you're better than some is not as satisfactory as the other feeling of not being on top is dissatisfaction.

*H*15 I wanted to make as high a score as possible because I want to stand as high in the group as possible.

*H*72 Wanted to excel others. "Sometimes I'm not even satisfied to be better than the best."

*H*38 I want to be above average. . . . I am above average. . . . I wanted to make as high a score as possible to show myself up to you—to show how high I could go.

*L*77 Just human nature to do as well as he can in any tasks whether competing with others or against yourself. . . . I suppose you might call it the desire for prestige. But I wouldn't get any pleasure out of being pointed out, having everyone say, "There goes the great so-and-so." I wouldn't get any particular kick out of that. [Freud might think he would.] But in these tests I would like to do better than Joe Jones.

*L*42 In the back of my mind I was thinking of what other fellows who've been here had done. If you had told me 80 had been done on the target I probably would have shot for it. . . . [Working proficiency distracted by] thinking of the score, i.e., to see how high I could reach, to see how much I could get done.

*L*64 . . . Naturally one wants to do as well as one possibly can on these things. . . . Everything you do you want to do well in. . . . I imagine there's a certain amount of vanity, ego, urge to succeed, to do well, to be superior. . . . I don't know why. I just intensely dislike seeing people do things very much better than I, especially if they aren't older. As a matter of fact there are some things I don't like anybody doing better whether older or not.

*M*69 I suppose I wanted to do as well as I could. . . . when there are several people taking the same test I like to see if I can't do better than they. . . . It seems to bolster my opinion of myself.

*M*62 Subject claims he would only obtain complete satisfaction in this experiment if he were "on top amongst the first few. . . . I always try hard and I like to see my efforts aren't wasted. . . ."

*M*51 . . . I'm not the only one taking this and you always try to get yourself near the top. That's natural. . . . You always try to succeed—win out. . . . You know there are a lot of people taking the tests and the odds are 100 to 1 against you and you pull the odds and want to come out on top.

*M*37 My idea was to get the highest scores possible. . . . I suppose I want to be considered a genius.

M32 I regard this as a competitive situation. Whatever activity I'm in I like to be first whenever possible. Of course here only you would know whether I was first or not—even so, it did spur me on. . . .

(*b*) *Wanted to do better each time.*

H25 It was sort of amusing to see whether I could do better . . . competitive spirit.

H18 Wanted to do better each time but didn't always expect to. "I always like to do better in anything I do."

H17 I always have a strong drive to do better—I always try . . . except when you're lazy but in this you can't be lazy.

H29 Wanted to do better each time "Because I didn't want to appear dumb I guess. But (n.b. immediate cover-up) I'm not particularly worried whether people think I'm dumb or not.

H35 Wanted to do better each time, to do at least as much as estimated and preferably more than that.—"To flatter my ego I suppose, to make me feel good—how great I am."

L63 After all, you're supposed to try. I mean you have to do better. . . . You have something within yourself that tells you you have to do better, or at least try to—perhaps ego—or could it be conceit? You can't show yourself to be too stupid.

L53 I like to do better and felt I should do better each trial.

L14 I think everybody would try to do better, although if anybody wanted to fool you they could just stop at what they estimated.

L47 There were times I found I wasn't [doing better] and then I'd just take note of that and then I would try a little harder. . . . Well I figured you're going through all this trouble I might as well do my best . . . maybe it's the way I was brought up. . . . Well they always say "You can do better than that." . . . When I bring home a B+ they say "Why didn't you get an A. . . ." Also "If a thing's worth doing at all it's worth doing well." Someone told me that once and it made a very great impression on me. On the whole would not say he was very much affected by his scores.

M57 I wanted to do well in them—much better each time than I had done previously because I thought I was capable of doing better than I did and I wanted to reach a performance I thought I was capable of reaching. . . . Another reason I suppose, consciously or otherwise, I wanted to beat the scores that anyone else had made. . . . Everybody has a certain

> opinion of himself, of superiority and standard of himself, so that he wants to beat others in order to substantiate that opinion so that his ego can face reality. (N.b. He had recently heard a lecture on Adler and was influenced by his terminology.)
>
> *M73* I wanted to do better . . . because I think these would be compared with other boys' grades and although it doesn't have any effect on my grades in school I believe everybody wants to rank as high as possible.
>
> *M61* Improvement is always good for the system. It always gives you more confidence when you make progress, whereas otherwise you're in doubt. And of course I always thought, "Someone else did better" and there was that incentive. . . . I thought the harder I tried the more chance I had of coming up among the leaders. . . . I've always felt you had to be pretty far up in order to get any place. . . .

Other bases for "trying to do better" are revealed by the following reasons: "in order that my opinion of myself would improve for that particular test"; "to give you a feeling you've done pretty well. . . . I like to be successful"; "I just don't want to be stupid. . . ."; "I just think I'm good. . . . I like to think myself good enough to do better each time I try anything"; "an average person should learn a little. . . . I should hate to be subnormal—I mean I should hate to know it"; "largely because a low score shows inferiority and if I am inferior I don't like to show it"; "it might possibly be (*sic*) that I thought I might be compared with a standard which I wanted to equal or better"; "to show off I guess. . . . I imagine other boys have taken these tests and . . . I would like to . . . see how much I beat them or they beat me"; "well, I enjoy showing myself how smart I am, that I can learn and improve with practice"; "another reason might (!) be that I didn't want to appear to have too bad a score"; "you can't show yourself to be too stupid"; "to save my face"; "because I didn't want to appear dumb"; "to make me feel good—how great I am."

The extraordinary awareness of competition, the large number of subjects who wanted to be on top, or to get the highest score of all, and who wanted to do better, in fact even felt compelled to do better under conditions which were competitive only by virtue of their attitudes, reveal the strength of cultural factors clearly. One notes again and again the repetition of the stereotype—allied to the illusion of vertical mobility— that improvement alone was a mark

TABLE 13

Goal-strivings in the High (*H*), (*M*), and Low (*L*) Median Average Difference Score Groups

Group	To be on or near the top		To do better each time		To do as much as estimated		To do as much as he could		Miscellaneous
	No.	%	No.	%	No.	%	No.	%	
H (*N* = 19)	10	53	16	84	1	5	—	—	1 volunteered he did not always expect to do better.
M (*N* = 34)	17	50	27	79	1	3	3	9	1 "just sort of did my best and let it go at that."
L (*N* = 18)	8	44	13	72	3	17	1	6	

Information about the certainty of one's expectations to accomplish or not to accomplish one's estimate or actual level of strivings may be significant, but unfortunately it was not checked on during the experiment.

Some subjects contributed more than one reason to the question, "Did you always try to do better? . . . Why?"

of intelligence, ability, and so on. The clearest presentation of this attitude was made by *M*50:

> Well naturally you always want to do better the next time. . . . to improve yourself, and get ahead of the next fellow I suppose. . . . I guess it's just the idea you get after a while—everyone trying to get ahead of everyone else. . . . [I wanted] to get on the top of the heap I suppose. I guess it's the idea everybody gets ground into them. . . . I want to be near the top.

One must not overlook, however, that all this material is on a verbal level. We have no information whatsoever on the intensity of the energy directed towards achieving the verbally desired level of attainment. It is probable that not all of them feel the same intensity of desire or need to be "on top," or to be "above average." If this is true then real differences in goal-strivings may exist in the three groups despite the similarity in verbally expressed goal-desires. Individuals probably differ with respect to the extent with which cultural standards and symbols and goals are internalized, so that though two persons may express themselves in exactly the same way, one exerts greater energy to achieve the goal because it has become an integral part of his personality structure and "needs."

V. ANALYSIS OF INTERVIEW MATERIAL (Continued)

A. Feelings of "Failure" and "Success"

Since the experimental situation here is relatively unstructured, i.e., the individual does not know the meaning of his performance scores, and since most individuals are highly motivated in the sense that they want at least to compare favorably with other subjects, cultural definitions of failure and success, as well as individual definitions which are partially dependent on cultural standards, can operate to influence the individual's reactions to his scores. The most pertinent cultural standard, in this connection, has already been referred to in the preceding section, i.e., the experience of success that is attendant on "improvement" and definitely absent on "decrement." A feeling of satisfaction, if not success, was often experienced even though the individual ostensibly failed to achieve what he set out to, since cultural approval is extended to those who "try to do their best," though they do not succeed.

Hoppe (7, p. 22-24) has given a careful analysis of the differential effects of failure and success on the individual and his aspiration level, pointing out that not all successes and failures have the same importance, that one can distinguish different degrees, namely, "full," "partial," and "simple," and that a rise in aspiration level most often follows full success, while aspiration level falls only after a series of failures. "Single failures as a rule had no influence on the movement of aspiration level, especially if it occurs in a series of successes." When Hoppe refers to aspiration level, he means the actual level of momentary strivings which he assumed was spontaneously revealed by the subject. It will be recalled that no estimates were asked for in his experiment and aspiration level was determined on the basis of various impressionistic criteria. The explicit estimates responded somewhat similarly to failure and success as the "aspiration level" in Hoppe's study. Individual differences are very much in evidence, however, subjects differing markedly in the immediacy with which they responded to increases and decreases in performance scores. It is true that in most cases a cumulative effect of failure is observed, i.e., a lowering of the estimate occurs only after a "low" performance is obtained two or three times in succession, or within a short time interval. But this cumulative effect, possibly to a lesser extent, may also be observed in connection with overt responses to success. A "full" feeling of success will not inevitably make for a change in the actual level of strivings. The actual level of strivings may not alter after failure either, though the estimates are decreased in accordance with the lowered performance. A sharp drop in score to one individual is considered a "fluke," something caused not by his own inability but by various and sundry external conditions, and his estimate is not changed. To another, an abrupt increase in score is considered accidental, while the decreases, howsoever abrupt, are responded to immediately with a decrease in the estimate. Some respond rapidly, at least in their estimates, to each new performance score. In most cases, we do not have knowledge of the changes in the actual level of strivings, following failures and successes,[13] when the level was not coincident with the estimates. We do know, however, that individuals displayed differential resistance to change following success and failure, as well as different degrees of resistance to change in general. An intra-test analysis may be one way of obtaining a measure of such resistances, which measure may provide a significant clue to, and basis of prediction of, the individual's behavior.

It is clear that neither success nor failure can be defined in terms of the absolute or relative achievement scores. To one individual 40 is a good score because that was the highest score he had achieved, and it had been obtained with difficulty. To another, a score of 60 in the same test is not very good, though he worked at top speed for it, simply because he had an idea that 75 would be a good score. It is possible that less individual definition of what constituted a "good" score would be obtained if the "average" score were known. That such knowledge would by no means make for unanimity is clear, however, since individuals differ in the degree to which they are susceptible to, or accept group norms. We would always find a few perfectionists who would only consider "100" a good score, though that were never achieved, as well as a few whose standards were completely responsive to the group in which they found themselves. As Hoppe pointed out, further, even for the same individual a given score will change in meaning depending on the course of his performance scores. If at the beginning of the task he thought 40 would be the highest score he could reach, when he reached that he would get "full" success, but after he passed it, 40 no longer would be held in such high esteem. In short, we are dealing with a variety of "frames of reference" which depend on a variety of factors that are not necessarily constant and will therefore make for shifts in reference, and perforce in the values assigned to particular scores. Feelings of failure are reported most frequently in connection with the three tasks having known proficiency limits—the synonym, steadiness, and target tests. Thus, too, feelings of success are reported whenever a zero score was obtained or approached in the steadiness test, and in the target test, whenever a bull's eye was attained. One does not find feelings of success in connection with the synonym test because the highest number of words ever recorded was 18 out of a possible 33.

---

[13] When the inner goal is not lowered simultaneously with a decrease in explicit estimate, the disparity provides a source of conflict, as well as incentive at times. Similarly the disparity between performance and estimate may arouse conflict or it may be regarded as an incentive.

Success and failure are only relative experiences to some; others have absolute standards for success and relative for failure, while still others have relative standards for success and absolute for failure. E.g., some never experienced failure in connection with school work unless they actually failed, but did experience success if their grade was better than the last time, or if they beat Joe Jones, or if they had simply answered a question their friends had difficulty with, although their total grade did not differ from that of their friends. On the other hand, some experienced failure if their grade was lower than the previous time, even though on absolute standards the lower grade was still in the "excellent" category, and some experienced disappointment, if not failure, when many more besides themselves attained an *A*. For a really complete description, of course, one must, like Hoppe, differentiate among degrees of success and failure under different conditions. It would be extremely interesting to know whether those who claim they are only satisfied with perfection feel it necessary for their "peace of mind" to attain perfection. In some cases we would suspect that the perfection-ideal is but one method of providing oneself with a "rational" goad, some would say symptomatic of a masochistic personality organization. Such individuals are never "satisfied" with any of their achievements. What the underlying basis for such reactions is, we are not prepared to say.

As an incidental example of the different reactions to the same objective fact, we find that the disparity between performance and estimate in some cases provides an incentive, and in some cases is a definite deterrent, avoided whenever possible. Thus we find some in the *H* group who claim to set estimates purposely which they do not expect to reach or may not even want to reach, on the assumption that the difference between their performance and estimate will help them get a higher performance score than otherwise. In the *L* group we find some who assiduously avoid setting an estimate which may not be at least equaled by their performance because any disparity of that kind would be too discouraging and keep them from doing their best. Only in rare instances, however, are disparities, where the performance is higher than the estimate, regarded as failure. Such disparities are necessary to some and unnecessary to others. This is one of the few clear differences between the *L* and the *H* group, i.e., the *L* group cannot allow the performance to be below the estimate, whereas the *H* group can. In the *M* group, on the other hand, we find a mixture of attitudes. *M57* remarks, "I thought that if I left a wider margin between the performance and the estimate, when I went above the estimate, I could get a feeling of success." There were some in the *L* group who felt the same way, and some who perhaps had greater awareness of what they were doing, did not experience success when the performance was above the estimate because they considered the estimate "low" to begin with. And finally, the size of the disparity adequate to provide an incentive or to act as a deterrent varies with different individuals.

Although we have insufficient evidence on this point, there does not appear to be any inverse relationship between the frequency of feelings of failure and the frequency of feelings of success. There are individuals, of course, who cannot view any of their activities in the light of a failure, and individuals who find it difficult to regard anything they accomplish as success. At the extremes of responsiveness and unresponsiveness to external events, we find many and very few experiences of failure and success respectively. There is some evidence in the interview material to indicate that no subject experiencing little failure in this situation experienced it frequently in life, and vice versa. This is not surprising since the problem of frequency of experiences of failure and success is probably dependent on what may be called thresholds of resistance to particular events, and since the individual is responding to the experimental situation, his particular thresholds would thereby be revealed.

A brief summary of the different causes for failure and success is relevant at this point. One will note that it is the reactions to failure rather than the causes of failure, that throw the most light on the problem of the genesis of changes in estimation.

The most frequently mentioned causes of feelings of failure were that the obtained performance scores were "poor," "low," that the performance score had decreased rather than increased, and that the performance was below the estimate. This numerical account (Table 14), does not tell us very much, since it is not the number of tasks in which failure was experienced, but the intensity of the reaction to that feeling, that is of real significance.

TABLE 14

CAUSES OF FEELINGS OF FAILURE IN THE THREE MEDIAN AVERAGE DIFFERENCE SCORE GROUPS

Group	Performance "poor"—"low"[1]		Performance had decreased not increased		Performance was below the estimate		Disappointed in all of the scores	
	No.	%	No.	%	No.	%	No.	%
H (N = 18)	10	56	2	11	6	33	2	11
L (N = 19)	16	84[2]	3	16	7	37	2	10
M (N = 34)	15	44	7	20	6	18	0	

[1] This category includes reports of feelings of failure arising out of scores *obtained in part of a task or the whole task. The steadiness test was* mentioned specifically 13 times, the synonym task 23 times, the symbol-digit task 8 times, and the target 11 times, in this connection.

[2] The critical ratios of the per cent frequency in this category between groups H and L, and M and L, are less than 3.

It is interesting, however, that a number of subjects in each group rarely or never experienced failure, specifically 6 in the H group, 5 in the L group, and 9 in the M group. The most frequent reason for not experiencing failure is based on the phenomenon of "weighting of the future" referred to previously. The following is a sample of the kinds of responses obtained.

Rarely or never experienced Failure.

*H*23 Practically no feelings of failure since concentrated on "doing better next time."

H25 No feeling of failure when performance below estimate since usually felt he could do better next time.

H18 Never really had feelings of failure because "knew I had another trial" and thought he would do better.

H35 Was discouraged by performance below estimate but felt he would do better next time.

H19 Had but few feelings of failure because he kept shifting his attitude towards the estimate. Only if his performance was continually below estimate to a marked degree would he experience failure, and adjust the estimate downwards. Under all other conditions he felt he would do better the next time. His attitude towards the estimate, however, depended on the actual outcome of his prediction. If performance was below the estimate he considered the latter as a goal which he really did not expect to reach; if the performance was equal or nearly so to the estimate, he regarded the latter as an "honest expectation."

L52 No feeling of failure when performance below estimate because "I figured I could make a comeback. . . . There's no such thing as a total failure. You can always make a comback."

L48 "I just don't let it worry me. I'll try next time." This is exactly his attitude towards marks in school too.

L53 No feelings of failure in this situation, or in school or social life. We have already discussed the protective nature of his conscious level of strivings and here we learn of its effectiveness. In this situation, since his performance was rarely below his estimate and since "only when I fall below what I aim at it bothers me," both the reason for the nature of his estimates and the reason for the absence of failure experiences are disclosed.

L60 Actually rarely experiences failure under any circumstances because of his highly developed rationalization system. When called upon to predict any outcome, he attempts to do so in a way that will take account of all possible, even contradictory, events. He is discussed in detail below.

M73 Rarely experienced failure except for very temporary flashes, since readjusts himself almost immediately to whatever height of performance he actually achieves. Such adjustment is based on a very simple and direct philosophy: "Just accept whatever you are."

M69 I never experienced failure because I always felt as though I was doing average or better than average.

> *M58* "No feeling of failure—maybe at the beginning with words. Then I got suspicious that many of them didn't have any answers at all. . . . They did? Then they were so far beyond me it didn't make any difference. . . ." This is an interesting confirmation of Hoppe's finding that too difficult material does not give rise to feelings of failure—which can best be understood perhaps, in terms of Lewin's concept of "barriers" (10).
>
> *M59* No . . . I didn't feel any failure. I felt I could do better but I wasn't dissatisfied with it.

Reactions when the performance was below the preceding estimate varied from feelings of failure, disappointment, discouragement, and "weighting of the future," to combinations of these. Not infrequently, one finds feelings of disappointment accompanied by a feeling that next time will be better, and occasionally disappointment accompanied by discouragement is reported, a combination which, unlike the former, tends to disintegrate the individual's efforts. It is extremely interesting that a decrease in estimate, following a performance below the preceding estimate, was made by both individuals who, as a consequence of the decreased performance score, were spurred on to work harder and expected that next time would be better, and by persons who were discouraged by the drop in score and weren't sure they could do very much the next time. Individuals differ, moreover, in the number of points below the estimate which would give rise to the reactions to failure indicated below, depending on the meaning of the estimate to them, the instantaneousness of their rationalizations, and the general intensity of their emotional reactions. Thus it is that an account of the number of people who reacted in a particular way in each of the three groups would not be of much value without the additional information on the size of the disparity necessary to call forth the given reaction.[14] It might also be important to find out whether those who are more resistant to lowering their explicit estimates are also more resistant to lowering their level of momentary and ultimate goal-strivings. We would expect no direct relationship except in cases where one could validly speak of a "generalized" stubborness under stress.

---

[14] An intra-test analysis would be helpful in this connection. Many subjects affirm that they "didn't cut down in effort" when performance dropped, but since feelings of discouragement and disappointment were accompanied by desire and determination to do better the next time, no doubt discouragement had differential effects on the strength of application of different persons. Such information may be irrelevant to our immediate problem, but since the desire to do better and the effort expended in doing so determine the future performance and subsequent levels of momentary strivings, aside from the explicit estimates, it may be important to know about them.

The causes for feelings of success are naturally the opposite of those which aroused the feelings of failure, namely, "high" performance scores, improvement of performance, performance scores above estimate, and occasionally performance scores equal to estimate. It is interesting that there are a few in each of the three groups who experienced no feelings of success at all. For comparative purposes we quote the reasons offered for not experiencing success feelings.

*H*13 Apparently did not allow himself to experience either success or failure very distinctly in this situation. Success was not attendant on performance above estimate. "I felt pleased with myself but never felt that was the best I could do."

*H*38 Did not experience success when performance above estimate "because I was looking forward to the next time when I'd fall below what I said. I didn't feel I had mastery over the test which gives feeling of success, so that you know what you'll do next time without having chance play a part." Regarded himself a "failure" in the word and steadiness tests. His definition of success is extremely significant. The physical handicap from which he suffers has made him particularly sensitive to situations and necessitated a severely planned life from which chance had to be excluded as far as possible.

*H*25 No feelings of success even when performance above estimate. "You felt you were capable of doing that regularly and it was just a fulfillment of it."

*H*72 Only experienced success to a certain extent when performance equal to or above estimate, since "on the whole . . . I always wanted to reach higher" than any obtained performance score.

*L*64 When asked whether he had experienced success at any time during the experiment, he replied, "I couldn't tell you that until I heard what others had done."

*L*78 Would have feelings of success only if told his scores were above average. Did not experience success when performance above estimate because "Possibly (*sic*) I had a feeling I gave a low estimate simply so that I could do better so that it didn't give me a feeling of success at all."

L31 No particular feelings of success when performance above estimate—"only by comparison with feelings when performance was below estimate." "High" scores did not elate him either because "I considered the task was kind of simple and I should have done that all along."

M80 Claims that he has no feelings of failure or success in life and had none here.

M67 No real success—just transitory, "for example, when you just fly through on the cancellation. . . ."

B. REACTIONS TO FEELINGS OF FAILURE

Reactions to feelings of failure, as is known, differ widely among individuals and in the same individual, depending upon the intensity of the feeling, the importance of the thing "failed" in and the conditions under which the "failure" occurred, i.e., whether he was alone or in a group and if in a group whether among friends or foes and so on. This does not mean that behavior is "specific" to the situation and that prediction of behavior is impossible. The "different" behaviors may be considered as variations on a theme. Where the theme is known the variations are seen as integral parts and congruent with the source. The theme in the individual would be his reaction to failure "in general" at any given period of his development. Thus we could say that *A* "withdraws," *B* is spurred on to work harder, *C* just "gives up," *D* loses interest and turns to another field of activity, *E* rationalizes, and so on. Within these broad categories finer distinctions could be made, e.g., B_1 is spurred on and is not too upset to accomplish what he sets out to do, while B_2 is spurred on temporarily but is too disintegrated by anxiety feelings to achieve his aim and finally "withdraws" by turning to something else. Or, A_1 "withdraws" by escaping from reality and indulging in autistic thinking, A_2 "withdraws" by repressing feelings of failure, losing interest in the external world and becoming prey to anxiety feelings which incapacitate him for further work until some external stimulus proves sufficiently strong to bestir him again, and A_3 withdraws physically from the situation, feelings of failure and anxiety feelings are both repressed, and he reacts with aggressive hostility towards the situation and all persons connected with it. Murphy, Murphy and Newcomb (13, p. 214) point out that "failure to achieve a goal does not leave the organism either quiescent or subdued—the organism turns restlessly to new efforts or turns its attention to new values." It is important to note, however, that failure does subdue some individuals, for a greater or lesser period of time, while others can immediately rationalize their failure and release the tensions thereby.

We can speak of generality as opposed to specificity of behavior if we understand generality in terms of the meaning and value to the individual of the situations to which the individual is responding. There is no reason to call a person inconsistent, or lacking in generality of behavior, if he responds differently to events which have different meanings for him, but which may appear objectively similar to an observer. A particular course of events in one task may not mean the same as in another simply because the tasks themselves do not have similar values. Under such conditions we need not expect similar reaction-patterns. Or the standards of performance and actual performances may differ more widely in one task than another, eliciting feelings of failure in one and not in the other. Why should we expect similar difference scores under such circumstances? Further, why should one's reactions to a very intense and very mild feeling of failure be the same? The need of protecting oneself, for one thing, is less intense when a mild feeling of failure is experienced. To some persons, of course, there are no degrees. Failure and success are all-or-none experiences. But such an attitude is not common.

Since we have no way of measuring the intensity of the reactions to failure, the quantitative survey of the types of reactions occurring in each of the three groups will not be of as much value as a sample of the responses reported below.

(a) Spurs him on to try to do better.

*H*23 I think that falling behind spurs me on more—gives me a feeling of insecurity and I made a greater attempt to obtain a better score. [When performance below estimate, he thought]: Well if I didn't do it that time I'm going to try hard and do it next time.

*H*28 Spurs him on up to a certain point only, then, "I just let it go."

*H*65 "I've never felt anything really had me licked. If I failed it once, with more effort I could do it well enough to get by at least." At the same time feels very despondent and becomes panicky. "If I didn't do as well as I did before

I was inclined to go wild. If I hit 30 and estimated 35 and hit 28 I would probably go back to 35 or even a little higher. . . . Just that I had failed and was annoyed with myself and determined to overcome the losses—but if I did better than the previous performance I was inclined to go higher but not to any wild extent."

H45 If you fail there's a stronger urge to be successful. I think you have a tendency to remember successes and try to forget failures and failures are often compensated, i.e., you try to overcome them.

H19 I try to do better but it has a discouraging effect.

L54 Apparently tries to do better, but not with much zest. "It is not necessarily that I give up in case of anything I fail, just that success is more encouraging." Feeling of failure more intense than feeling of success.

L16 Failure spurs him on to do better more than success—"because of rebound to having gone below." He immediately lowered his estimate, however, when performance was below estimate, though neither goal nor effort seemed to decrease.

L31 Had extreme reactions to feeling of failure, either lowering or increasing the estimate drastically, doing the latter "in order to force myself to do better."

L53 Failure always bothers me a lot and I try to do more about it. . . . Mostly I see where I could do better.

L52 Only at times did failure spur on; there was "a thought of scoring higher and coming back and some of them when there was a low score there was a loss of interest."

M6 When experienced failure he "put forth intense effort and worked like anything."

M69 "Failure would naturally spur me on more. . . . For the simple reason that it was a failure and I don't think I should fail in doing anything. But I don't think success spurs me on." In answer to *E*'s question concerning the reason for the relatively high estimates in the target and s-d, he replied, "In the case of the game I felt I should be getting scores over 60 and there if any place I had a real sense of failure when I went below, so my sense of failure there tended to spur me on more and more to do better and made me raise my estimates higher than I might have ordinarily. . . . When I did make a low score and then made a high estimate it was sort of added inspiration to do better."

M49 "Usually failure is more intense because you take the success more or less for granted, whereas you take steps to avoid a repetition of the failure and since you do try to avoid it, when it does happen it makes more of an impression. . . . Both failure and success spur on. If it's failure in a thing you have no respect for, it does not bother you, but if it's a failure in something you respect—I keep plugging at it. And success naturally spurs you on in anything since you like anything you're successful in and you feel sure you can continue to be successful." There are wide individual differences in the assurance regarding continuation of a success.

M26 "Neither failure nor success more intense either in this situation or outside. I think they're just about even. . . . Failure tends to spur me on more. . . . I don't know why. It just seems to. When I hit it, it's sort of relaxing and I don't work so hard on it." This kind of reaction to success may be typical of, and symptomatic of anxiety feelings in regard to one's ability.

(*b*) *Sedative effect; discouraging; disorganizing.*[15]

H72 Failure usually has a "sedative effect," although occasionally gets very mad and tries to do better, "but not usually." "I've had failure so much. I've failed so often that it doesn't inspire me anymore. I've passed that point. Whereas if I achieve something and it works out, well that justifies faith in myself. A certain amount of failure is good but in my case where a person has failed so much it's inclined to act the other way, sort of makes you lose faith in yourself."

H38 Gets discouraged but tries to console himself by thinking of all those who got even lower marks than himself.

H44 Does not admit it, but behavior throughout shows at least temporary upset and disorganization when he did not do as well as expected. After a while he would just lose interest, and feelings of failure would be allayed.

H33 "Reaction is to give up—but I don't do it—I haven't done it. . . . There are times when reaction is to work harder—that is, when you see that the errors were superficial which had given you a C instead of an A. . . . Success definitely spurs me on. . . . Failure is a more intense experience. I feel that it might be a hindrance in a social way. I wouldn't like to have it known amongst my friends."

---

[15] Each of these effects are qualitatively distinct, but they are grouped together since they all represent what mental hygienists would term "unhealthy" reactions to failure. All three, moreover, result in a greater or lesser cessation of directed activity towards the original goal.

L9 "I am depressed by failure simply because it robbed me of the feeling of success. When I get a failure I don't resign myself but I am depressed." Also tends to lose interest in the offending task or field.

L77 "It depends . . . sometimes I get sore and bust around the room and want to break everything. . . . At other times I'm just discouraged. As a general rule it doesn't tend to spur me."

L78 "Failure makes me moody. Rather than a spur, it's probably an anchor." Experience of success is more intense and spurs him on. "I probably contrive an excuse for the failure and chalk it up to experience. Success I always think it's my own doing."

L64 "I get awfully mad when I've failed something. I was so awfully mad at my drawing teacher since he had no right flunking me that I had to go out of the room to keep from throwing things at him. . . . *I hate failing.* If I fail in something I know is my own fault I'm fit to be tied and if it's someone else's fault it's even worse." His energies seem more engaged in "hating" than in overcoming the failure.

M20 "For a while failure makes me want to quit, to stop trying, but the more I think about it, the more I want to try. Also it depends on the failure. If it's just slight failure I feel, 'Well, the next time I'll do better.'"

M61 "I can't recall ever having any definite feelings of success. Success seems a little more temporary. . . . Failure is more lasting. After you do have a success you go right on and set another goal . . . whereas with failure you mope about it a little while and just wait until something else comes along. . . . There have been more failures than successes. . . . Neither failure nor success spur me on to do better—just falling a little below the point of contentment, a little below what I know I can do. That's where spurring comes. If I am a failure I just don't overwork and if I am a success I just relax. . . . I take failure as a matter of occurrence in some cases, like social life, where it becomes anticipated, and in some cases it's a disappointment."

M82 ". . . a very depressed feeling whether it occurred in public or alone. Of course you rarely actually have a feeling of failure in front of people because nowadays they just joke about it. They just say off-hand, 'I guess I'll never make a good math student.' . . . I go rowing and if I'm not in the first boat I'd feel very poor. . . . In general I keep my ego pretty well up. I don't let the failure get the better of me. It's quite intense for a while but then I turn to other fields."

(c) *Feeling of failure of short duration and relatively slight effect.*

H76 "Feelings of success more intense. I always used to like to compensate for feelings of failure by things I'd been successful in, in my own mind, and naturally the times I felt successful I didn't think of the times I had failed. It's simply that I don't let myself feel I'm a complete failure and just thought of the times I hadn't failed . . . and felt they at least took the place of the times I had failed, if not superseded them in extent and therefore I like to inflate the successful things in my own mind." Actually appears to repress own anxiety caused by feelings of failure and to project it on to his parents and friends. "Failure usually affects me in worrying about the possible disappointment of my parents or my friends who might have the idea that I have ability and might get someplace."

H19 "In times of failure I was consoled by previous success in another preceding task. I felt, 'Well I had just slipped here and I might get back.'" (This kind of reaction was also observed by Hoppe.)

H18 Rarely has feelings of failure both because always reacts by placing emphasis on the future, and because he makes a fairly rapid adjustment to circumstances. "I'm inclined to optimism in any sort of events. This overbalance of success is not because I've had more success than failures, but because anything has to be pretty heavily weighted to produce abject failure. . . . I've never had a lasting feeling of failure because failure connotes more or less the end. . . . (Had very high marks in high school.) In English finals I got a C. . . . I had a very intense feeling of failure . . . and now C's don't disturb me so much."

L70 Rarely has feelings of failure "because I don't want to think about it." On the other hand he has few feelings of success either because he does not think he has achieved anything of importance, and what good has happened occurred without any impetus from him. "They just affect me." His reaction to performance scores below estimate was "probably to reduce the estimate—together with some push to do better."

L52 "No sense crying over spilt milk, that is, if you failed you failed. . . . In failure there's no such thing as a total failure. You can always make a comeback."

L48 Feels depressed momentarily, and then tends to forget about it as quickly as possible by thinking of "successes." The depression does not last long because "Well I tried and that's all there is. . . . What can you do about it?"

L47 "If I don't enjoy something I forget about it in a few days. . . . (Neither feelings of failure nor success are very intense.) In general I don't think I ever think of success or failure. . . . I always feel, 'Gee, I could have done better,' and 'Well I'll do better next year,' and at the end of the year I say, 'Well you'll do better next year.'" It is not clear whether he actually exerts much effort to improve. His philosophy is distinctly of a retiring sort. The only clearly "spurring" factor seems to arise chiefly because his mother and sister express dissatisfaction with whatever grade he gets that is lower than an A, for example, "If you got a B you could have gotten a B+," etc. His expressed desire to do better may be simply an attempt to avoid the unpleasant reactions of the family.

M6 In many things he quickly forgets about a failure, "since it is past," but personal criticism, which gives him a feeling of failure, "bowls" him over, the hurt often lasting for a week.

M59 Contends that he has no real feelings of failure in school or social life. Even though he has flunked French three times, he does not feel a failure. "I guess I'm no linguist. . . . Nope, doesn't worry me in the least. . . . I just feel as though I never could get those languages and even though I am failing, it doesn't worry me because I just don't care about them—not interested." The rare times he does experience failure, "I think about them for a while and then just forget them. I pay no attention to them. . . . Yes, probably failure is a more intense feeling. . . . You feel as though you lost something that you really wanted and it makes you feel bad, and then you start thinking about something else to take its place and you start feeling better again. . . . I just don't ever feel I'm a failure in things—not for any length of time anyway."

These quotations give some idea of the complexity of a "reaction to feeling of failure," of the different types of reactions to the same event, of the fact that the immediate and subsequent reaction to the same event may be different in the same individual, of the fact that one may be spurred on but discouraged and depressed, or spurred on without any accompanying conscious depression or anxiety, or neither spurred on nor discouraged, of the fact that some are consoled almost immediately, some turn to other fields, some hark back to past successes, and some simply will not recognize any event as a failure—" . . . *because I have faith in myself* . . . ,"—and finally of the fact that a failure in one thing is not necessarily reacted to in the same way as in another thing because the two do not have the same meaning for the individual. We find also that whereas slight failure may spur one individual on and "intense" failure completely narcotize him, another individual is spurred only when the failure is very acute, a "slight" failure leaving him indifferent and disinclined to continue work. These reactions to failure are undoubtedly expressive of some basic aspects of the individual's personality organization, a function of his past experiences, general "sensitivity" to the environment, and indicative not only of the presence or absence of anxiety feelings but also of the presence or absence of repression of such feelings, and the extent of the field of his insecurity feelings. Undoubtedly aggression, insecurity, aspiration, and achievement are closely related to one another and to the ego structure of the individual. Only by depth-analysis, however, shall we be able to discover the ramifications and interrelationships existing among insecurity, anxiety, hostility, aggression, goal-strivings, ability to handle given materials (cf. 9, 16) and reactions to feelings of failure.

The most pressing questions at present are: What techniques are utilized to avoid failure? How are these techniques developed? What techniques are provided for, or rather sanctioned by a given culture? Can individuals be reliably differentiated on the basis of their "failure techniques" so that their behavior in specified situations can be predicted?

These questions are important not only from the point of view of personality study in general, but also from the special aspect of personality under consideration here, the level of momentary strivings. This is so because probably the most basic need, common to most humans in every culture is the need to avoid failure since failure offers a threat to the most cherished value and object, the self, howsoever the essence of self be conceived. Failure must be avoided because of the potential catastrophic resultants, i.e., losing faith in the one force that makes for unity of being and acting. From early childhood on, one is impressed by the need to protect oneself from outside belittlement, possibly because disapproval meant the temporary loss of love and security, and in time this need becomes more or less internalized and it is also necessary to protect oneself from self-disapprobation. Experiences of failure would be the most severe taxing of a laboriously won psychic equilibrium.[10] The "level of aspiration" situations offer just such a threat and challenge to the ability and very integrity of the individual. It would seem doubtful then, whether in any situation fraught with potential failure one obtains an accurate measure of the individual's level of strivings. Rather, the average difference score would appear to be more nearly representative of a kind of protective mechanism against failure, than indicative of differences in the height of aspiration level; and, moreover, in certain cases at least, the general characteristics of this protective mechanism are similar to the individual's reactions to any life-situation which offers the same problem. Unfortunately we have little reportable data on this last point, but what material is available is corroborative.

> *H*28 It is characteristic of me to set my goal a little higher than I'm capable of doing, it seems. . . . Also my goal is not set at a given mark but within a range of 5 per cent or so.
>
> *H*72 . . . Everything or nothing—you've got to aim high . . . if you aim lower than what you think you ought to do or hope to do, or like to do, it does increase your chances for success but it isn't as satisfying as if you aimed high, it's not quite so satisfying—or maybe it is. . . .
>
> *H*33 Gives estimates higher than actually expects to do

---

[10] To do more than present this oversimplified picture of the reason why one of the basic needs of people is to avoid failure would entail a systematic theory of personality development and structure. Such a theory would require an explanation of the fact that some persons have a definite ambivalence towards this self-object, ambivalence which differs in degree and intensity, extending from a "normal" degree of sadism-masochism to the final stage of self-destructiveness—suicide. We are not prepared to offer such a systematic theory, although we suspect that Freud's concepts of super-ego and ego, with the possibility of these two systems differing radically in their values, might be the most fruitful approach to an explanation of ambivalence, of sadism-mascochism directed against the self, i.e., almost a need to punish the self-value, of the aggression directed outwards upon other objects, of insecurity, of the extent of the field in which the individual feels insecure apart from that wherein he feels secure, and so on.

since feels he can get highest possible performance score that way. "It seems to tend to bring more of my forces to concentrate than otherwise. Might be a tendency to say, 'That's good enough,' if I didn't put a higher goal.'

L70 In general tends to play "safe." Subject indulges in self-doubts with a definite tendency towards self-disparagement of a kind that might well be called "neurotic." Thus, concluded his statement regarding keen desire to write with, "Silly, isn't it?" Has a B plus average but steadfastly refuses to recognize that this is above average. Insists that difference between a C and B average is not one of ability and furthermore the average student at Columbia gets a B average. Doesn't think he's "pretty good" "simply because I don't believe that any proportion at all of the good incidents that have arrived have had anything to do with me except as affecting me." He resented the question, "Do you ever have feelings of failure?" because he felt I was trying to have him say he was pretty good. On the whole seems to accept, in a somewhat fatalistic, resigned fashion, whatever he achieves at any time as indicative of the top level of his achievement. Aside from occasional blushing when disturbed, his face and manner of speech are very colorless and "immobile." One feels great insecurity, anxiety, and repression in this boy, together with a tendency towards negativism.

L14 Estimates generally not as much as hopes to get, nor even as much as he expects, but a little lower than the latter in order not "to be shown up."

L78 "Yes . . . always the least. Yet I always think to myself that I'll do better than the answer I give them. . . . Well it's probably to avoid being thought of I might say as a braggart or something of that sort. I want to be sure that they feel that when I give an answer that, well, that they can feel that that is absolutely the worst that I will ever do. . . . I imagine that they will hold me higher in their esteem on that account plus the fact that when and if I do better than I told them I'd do that makes me feel better too." This double function of low estimates is clearly set forth in a number of cases, but the reason that "others will think better of you" is advanced for maintaining both extremely high and extremely low estimates of future performance.

We can see that low intercorrelations among different life-situations would probably be obtained too, though this again in no way proves or even implies complete specificity and unpredictability of behavior. If we advance our knowledge to the point of understanding these basic motivations, we might, by the use of broad categories of description, be able to predict that under such conditions X will react in a cautious manner, Y will react in an evasive manner, Z will withdraw from the situation, and so on. But since, as we have seen, there is more than one way of withdrawing from a situation, and since some persons develop certain sequences of behavior under certain stress conditions which in their course include cautiousness, evasiveness, and finally complete physical withdrawal, quantitative results of different situations may yield even zero correlations, despite the qualitative similarities which exist among the behaviors evinced in the different situations.

Since no attempt was made during the experiment specifically to ascertain the different "techniques" utilized by the subjects to avoid failure,[17] the material available is scanty. The problem, however, in view of our interpretation of the average difference score, is to indicate at least the different kinds of protective reactions, and point out how the same general intent, namely, to avoid failure, may be carried out in quite different ways. Throughout the ensuing discussion, it is necessary not to overlook the dangers inherent in defining ends too broadly, thus subsuming under such ends several goals which differ markedly enough to warrant distinction from each other. On the other hand, one must also avoid inferring that widely differing surface behaviors necessarily connote widely differing sources of motivation. The writer engaged in this task of describing the major protective mechanisms with acute awareness of the necessary superficiality of material obtained by analysis of surface behavior. A careful reading of the excerpts from the interviews already quoted will disclose the source of the material presented below, and reveal at the same time, the usual absence of any "single" technique. In a society shot through with psychological inconsistencies, many avenues of expression and repression are open to the individual, making it almost inevitable that very few "pure" types will develop. Thus it is that in children we find active and passive forms of response highly intercorrelated. But increasing age means increasing experience, not necessarily all varieties, but experiences limited by kind of family, socio-economic status, appearance, type of group-affiliation, etc., and in accordance with the organism's general level of physiological activity and the preponderance of certain types of experiences, the range of possible behaviors becomes limited. Depending on the acuteness of the situation and the consistency of his past experiences, one kind of response will be more easily elicited than another equally possible kind.

---

[17] We do not mean to imply that all the subjects *consciously* utilized a *technique,* although there were some who did. Degrees of awareness of one's reactions and of the reasons for them differed widely among the subjects, and there were some who offered considerable resistance to even a superficial analysis of their motivations. The proof of a reaction does not lie in the self-awareness of it.

It must be remembered that all classifications are somewhat arbitrary, tending to obscure both the fact that "different" types frequently cut across each other and the fact that since these are all mechanisms designed to "protect," in situations of sufficient intensity all of them may be utilized by any given individual. These mechanisms moreover, are not equally applicable in all situations where the individual's self-integrity is challenged, and from a long-range point of view some of these reactions may prove more deleterious in their effects than protective, i.e., one may become obsessed by the need to protect oneself to such an extent that accomplishment is impossible. Delusions of reference and persecution exemplify this stage.

It seems difficult, however, mental hygiene doctrines notwithstanding, to state definitely that one kind of reaction is more "healthy" than another. Healthy for whom and for what? From the point of view of effectiveness of protection against feelings of failure, the most drastic method involving the shrinkage of one's world would be the "healthiest." From the point of view of the least inhibiting or most facilitating of one's efforts, in adults at least, one would have to go back to the individual, the situation, and the intensity of the given reaction before being able to pass satisfactory judgment.

One may attempt to avoid failure before it is met, or attempt to assuage and allay the failure after it is experienced. Avoidance may be accomplished by adopting a philosophy of stoicism, or a highly differentiated system of interests and values, or completely esoteric interests, or an indifference to all culturally and externally given values, or escape into a purely autistic world which in extreme form is pathological. Feelings of failure may be assuaged by loss of interest in the offending field, by instantaneous repression or rationalization, by fatalism, physical withdrawal, wish-fulfillment, concentration on past successes or future achievements, or by aggression against others, which may take the form of active or sullen hostility, belittlement of others and the task. Practically all of these reactions represent some form of *withdrawing* from reality in the usual sense of that word. Why some persons tend to utilize certain reactions rather than others is a developmental problem which we are obviously not prepared to answer. Nor are we in a position, at present, to link certain types of mechanisms with particular personality characteristics.

C. TYPES OF PROTECTIVE MECHANISMS

In this experiment the avoidance of situations, feelings, etc., which might arouse anxiety feelings and lead to experiences of failure could be accomplished theoretically in any one of the three major ways: high, zero, and minus difference scores. The estimate, or the performance score, or just the relationship between these two might be focused on, although strictly speaking no individual was able to ignore any of these three completely.

1. *Emphasis on Relation between Performance and Estimate.* One could attempt to concentrate all attention on the relationship between performance and estimate, i.e., on the accuracy of the estimate. One could ignore thereby, the meaningfulness of both estimates and performances, and concentrate on trying to avoid any disparity between the two. If successful, the result would be zero difference scores. That almost despite oneself other factors enter in is clear from the fact that relatively few zero difference scores are attained, and that only a few of the subjects consistently worked just to "hit it on the head," one of these being in the *L* group. These subjects, moreover, could not escape involvement with the actual performance scores as is indicated by their working to "do better" most of the time. What such an attitude did accomplish, however, was the avoidance of commitment of one's desires to *E*. This may be a habitual reaction to the environment or one elicited only in situations in which the individual feels insecure. *L*64 points out that he did not know what *E* wanted from him so that the only thing he could do was to estimate as closely as possible

to his actual performances. On the other hand we discover that as a general rule he believes that one should not "delude oneself" about future achievements, that in fact it is better to err on the other side, i.e., estimate lower than expectations and strivings. He did not attempt to delude himself about the actual involvement with his scores, however, since he obviously worked at top speed in every task, even averring during the interview that the particular task made little difference in his efforts: "I would have worked just as hard if you had asked me to push a peanut with my nose."

Such an attitude may also be accompanied by a more or less complete denial from consciousness of one's interest and strivings. *M*21 remarked, "In the back of my mind I knew my ability was being tested but I didn't know how much that interested me." The fact that he did become "interested" in the tasks where his scores went steadily upwards and that he admitted to feelings of failure when he thought his scores were too "low" suggests that interest was not really absent. Nonchalance of this type towards performance scores might in fact be interpreted as a reaction against inner caring.

An attitude of hetero-competitiveness may also be present along with the stress on the accuracy of estimation. Thus *M*26 explained that his goal was to estimate close to his actual performance, "closer than anyone else." His background has made him avoid competition on a performance score basis and, possibly as a reaction against the emphasis on *A*'s in his school teacher family, he considers "*C* a perfectly respectable grade."

This general orientation to the situation allows decreases in estimates to be made with less emotional involvement since they are presumably dissociated from the individual's level of strivings and expectations, merely representing a number which is to be matched with another number. This kind of retreat reaction from potential emotional involvement, if developed into a philosophy, might eventually lead to a desensitization of oneself—so that not only would emotions be avoided but when they occurred it would be with little intensity. *M*21's reaction to a situation fraught with emotional intensity and belittlement by another is suggestive in this connection. For a while he had been violently aroused by his girl-friend's taunting, then decided he probably did not love her any way, "and now," he concluded triumphantly, "she can't get as much of a rise out of me." Some individuals rationalized this type of withdrawal-reaction by trying to impress *E* by their "objectivity" and so-called "scientific detachment."

2. *Emphasis on Performance.* Dissociation of level of strivings and expectations from the explicit estimates was also accomplished by concentrating on the actually achieved performances. The question of what others had done acted as a spur to increase one's scores, but since the "average" score was not known, one could base one's reactions almost completely on what has been called "the fallacy of movement." Just as "society" wards off feelings of failure by making change synonomous with progress, so individuals here could overlook marked disparities between their performance and estimate or between their performance and actual level of strivings, if the estimate and strivings were not coincidental, by taking refuge in the fact that the score was higher than the preceding performance, or at least higher than the initial score. As *M*50 explained, "Some things I realized [realization a function of actual performance scores] you couldn't go much above and if I improved one point, I felt that was pretty good." Thus the fact of "improvement" alone, from initial to final scores, was sufficient to assuage and sometimes to completely prevent the feelings of failure.

3. *Emphasis on Estimates where Estimates are High.* When the individual could obtain no satisfaction from his performance scores, perhaps because he did not know whether they were good and feared that they might be poor, or perhaps because the scores tended to decrease instead of increase, or perhaps because he had a desire to impress *E*, or hide from *E* his particular feelings, the estimates themselves could be concentrated on. The estimates could be manipulated, if not the performance scores, to assure some feelings of success and satisfaction to the individual. E.g., if the performance is considered low, substitute satisfaction could be obtained by "high" estimates. This at least demonstrates to *E* that one aims high and as one subject put it, you could always say, "Well I tried." Individuals who put estimates quite a bit above their performances could feel encouraged because of the backing afforded by one of our cultural stereotypes, namely, if you aim high you're likely to go higher than you would if you aimed low.[18] The "aims" in this

---

[18] This is true only insofar as the disparity between present achievement and aim does not prove too discouraging, as it does with some persons.

experiment were never really set beyond possible reach of the individual, or as Hoppe described this phenomenon, "one seriously can wish to do only the possible" (7, p. 58). But failure need not be experienced when the performance is below, simply because to some the estimate represented a "goal" which they felt they did not expect to reach anyway. Occasionally this is a fairly confirmed attitude of the individual; at other times we find the subject aware of the fact that he has rationalized his disappointment. *H*19 reported, "The mark I tried to get was pretty honestly what I thought I would get but I probably rationalized going below by saying estimate was a goal. But I hadn't set out considering it a goal."

When the scores give the individual a feeling of failure, on rare occasions this need for substitute satisfaction becomes acute enough so that estimates are raised far beyond what they would have been had failure not been experienced. *H*65 reported on such occasions that he was inclined to go "wild." "If I hit 30 and estimated 35 and hit 28 I would probably go back to the 35 or even a little higher. . . . Just that I had failed and was annoyed with myself and determined to overcome the losses." *L*31 expressed somewhat the same attitude. "Whenever I had the keenest feelings of disappointment I either lowered my score drastically or else made it very high . . . in order to force myself to do better." Some individuals in the *H* group who linked high estimates with feelings of self-confidence, gave high estimates in order to hide from *E* their lack of confidence.[10] To decrease one's estimate under such conditions would be to admit openly that one felt he would do worse next time, which would be tantamount to admitting failure and lack of confidence in one's ability. The maintenance of estimates under conditions of great disparity between estimates and performance was rationalized by one subject as indicating his "perseverance," which characteristic he decided was being tested. Feelings of success could thus be derived from the "goal" rather than the actual performance. Similarly, some raised their estimates when performance scores did

---

Furthermore, one could actually have high aims but not be able to commit oneself openly to a particular achievement, either to another person or even to oneself, because of the disastrous effects on the self if the achievement fell below the aim.

[10] See more complete discussion of confidence above (p. 59 ff.).

not increase, in effect to give themselves the illusion of improvement simply because "I hate the idea of feeling, 'Gee I'm not improving.'" The intensity of this reaction to "improvement" is probably a function of the strong conviction that the absence of it was an indisputable sign of stupidity, whereas any degree of improvement indicated "ability." *H29*, who was doubtful of his intellectual abilities and who wanted "not to appear dumb," maintained estimates far above his performance, and, there is reason to believe, above his actual level of striving. Thus, we could interpret some of the high estimates as reactions against an inner feeling of inadequacy, an attempt, in some cases, to hide from *E* conscious feelings of *inferiority*. The utilization of high estimates as substitute satisfaction for unacceptable performance and the maintenance of high goals as well as high performance levels, under these circumstances, could be regarded as an attempt to relieve inner feelings of inferiority.

In some cases feelings of inadequacy regarding actual achievement may be so acute as to make for continual substitution for reality of the planning of bigger and bigger future attainments. The following quotation from Tolstoy exemplifies one of the psychic situations that may be responsible for relatively great plus difference scores:

> I found myself . . . in the excited condition of one who has *already lost more in play than he has in his pocket, who fears* to settle his accounts and who in his despair keeps playing cards without hope of regaining them in order to prevent himself from coming to consciousness (10, p. 147).

To maintain such a shaky superstructure requires a philosophy of life which would bolster one's behavior with a "rational" foundation. *H72*, who achieved the highest average difference scores in the entire group, has attempted to turn the unfortunate situation wherein his aims and desires are widely disparate from his actual achievements and probably from actual abilities too, into a desirable state of affairs. The following pattern of reactions has repeated itself many times during his life at school: He loses interest almost immediately in the things in which he does not succeed and frantically turns to another field of activity, where the same thing has always occurred, i.e., very high goals, very mediocre achievement, feeling of failure, discouragement, and a turning to some other field.

Thus, his philosophy visualizes everyone as unsuccessful in comparison with the vastness of the "field." He feels also a necessity to justify his maintenance of high goals despite inferior performance, not being able to realize that such goals substitute for his actual scores.

> "I expect high . . . i.e., I always aim high." *E: And how do you feel when you don't reach it?* "I feel terribly disappointed. *E: Then why don't you try to aim lower?* "I don't know. Because well I don't think its worth striving for. If you aim lower than you think you ought to do or hope to do or like to do, it does increase your chances for success but it isn't as satisfying as if you aimed high. . . ." *E: Do you actually think you aim so high because you usually are so dissatisfied with what you actually accomplish?* "That might be a reason so as to compensate for the inferiority you might say. Then I wonder if it is just that. (N.b. the shying away from such a dangerous admission.) I still feel that if a fellow wants to really contribute something or discover something he's got to aim high, if he's good or not because there's so much to the subject as a whole. I don't know how to express it. Think in terms of the subject matter, rather than in terms of your relation to others who might be distinguished in the field. . . . I think I'd make the very worst of anything or the very best. Everything or nothing. Nothing partly done. Never be satisfied with what you achieved. Always strive for something more, even though you know that other people think you are good, which has never happened in my case, still you should know inside of you that you really haven't accomplished anything, that there's more to be done in relation to the field as a whole. Even if other people think you're good you know that thing could have happened the other way and you could have been the worst failure. Call it luck, chance, anything you want that made you turn out to be an authority." The terrific anxiety and insecurity under which he labors (which is responsible for the development of this particular philosophy and in his case for the inordinately high explicit estimates which one feels represents a level of achievement which he neither expects to reach or wants to reach, i.e., he would be satisfied with performance below his actual "aims" if the performance were only higher than what he usually accomplished), is vividly described by him, revealing clearly the sapping of psychic energy by anxiety over one's ability. He apparently had to restrain

> himself from giving estimates that would be obviously fantastic. "I felt myself frustrated you might say. In my mind I expected more, yet found myself as if bound, fettered physically and sometimes mentally. If I could break free maybe something might happen. A sort of weakness comes over you. Your brain becomes sort of paralyzed and dead and doesn't want to go on. It is a funny feeling when you know that you must go on, that it would be disaster if you didn't, yet you remain where you are. . . ."

Realizing all this complexity of motivation involved in "high" estimates, we cannot possibly describe the basis of high estimates, as Frank does, in terms of the "need to keep the level of aspiration as high as possible."

4. *Emphasis on Estimates Where Estimates are Low.* This same feeling of discouragement with achievement led some individuals to put estimates below the score they were quite sure of attaining, either as a policy throughout the six tasks, or on special occasions. When utilized as a more or less general reaction to this situation, we find the fear of failure expressed as follows. *M68*: "I guess the reason I kept guessing lower than I did was that I'd rather guess low and do better than guess high and not do so well." *L16*: "I tried to make it a habit not to expect too much so that I wouldn't be disappointed." *M61*: "I felt rather than set a high standard and fall below I'd rather set a low standard and try to go above it." It is extremely interesting that in the *H* group one finds rationalizations for behavior couched in exactly opposite terms, i.e., "I'd rather aim high and fall below than aim low and just make it." In concrete terms, one subject told *E* that he would rather aim for an *A* and get a *B*, than aim for a *B* and get a *B*, although he would be satisfied with the *B* in the first condition, undoubtedly because the very fact of aiming for an *A* carries with it some reflected glory. *H33*, who aimed high because "it helps more of my forces to concentrate" and who "wanted to get as high a score as possible," nevertheless, said when he went "too much below my estimate, even though previously I had gone above estimate [n.b., attempts at consolation even in reviewing the circumstances of the experiment, by harking back to a past success], there was a tendency to build up from lower performance." Thus we see that estimates may be decreased far beyond what they would have been, just as they are sometimes in-

ordinately increased, when failure is experienced. It has already been pointed out that a decrease in estimate did not necessarily indicate a lowering of the level of strivings, but represented the individual's attempt to allay feelings of failure. When *L*52 was temporarily discouraged, e.g., "I set a number I thought I'd be able to beat in a way to give myself encouragement . . . to show I was starting upwards again.[20]

5. *Withdrawing from the Field.* Feelings of failure are sometimes resolved by what Lewin calls "going out of the field" (10). In Hoppe's experiment a bodily withdrawal was possible inasmuch as the subjects were permitted to quit a task at any time.[21] In this experiment, though bodily behavior was more restricted, mental withdrawal was not prevented. When a "bodily going-out-of-the-field is impossible the tension within the plane of reality leads to an inner going-out-of-the-field, to a *transition* from the plane of reality into that of unreality. . . . The affective tension expresses itself within the plane of unreality by restless thinking" (10, p. 149). This statement by Lewin is admirably illustrated by *M*10 who reported, "After I wasn't succeeding in rational learning for a while and felt I wasn't using the right system—my mind started to wander." This is probably the result of experiencing failure where failure is intolerable to the organism. Other subjects reported similar restlessness when the pressure of the situation became too great. E.g., *M*81 reported, "You might be interested in the way that in the addition I felt I was tiring and got peeved and was ready to blow up when I decided it wasn't worth very much. . . ."

An "inner going-out-of-the-field" is also illustrated by the "loss of interest" which sometimes accompanied feelings of failure aroused by unacceptable performances. This reaction is not limited to any one of the three major groupings but is found in all of them. When *L*52 was discouraged temporarily by a decrease in performance, "there was a thought of scoring higher and coming back and some of them when there was a low score there was a loss of interest." When such "loss of interest" occurred it was easier to decrease estimates in accordance with the achievement, than it would have been if interest was consciously sustained. Very frequently this type of belittlement of the task, together with the decrease in the estimate resolved the tensions aroused by the feelings of failure so that succeeding performances were better and interest was again aroused in the task. *There was one very striking example of such* a course of events. The subject began the target with obvious interest and eagerness but after the first two trials his performance began to decrease and the tensions aroused by this turn of events served to maintain this downward trend. At about the fifth trial he attempted to prove that the discs were of differential smoothness, but that failing, at the beginning of the eighth trial he suddenly remarked, "You know I'm not interested in this." This "explanation" of his low scores released both the tensions and the feeling of failure, allowing him to perform unhampered. The succeeding performance score was the highest he had yet achieved.

---

[20] Such temporary attitudes and behavior occurring unevenly in different tasks would inevitably lead to low average difference intercorrelations.

[21] In Hoppe's experiment a subject would occasionally quit after experiencing success. This may be attributed to the fact that he was not confident that the success was genuine and feared that he could not repeat the same performance. Analogous behavior was observed in this experiment when a subject would not increase his estimate, in fact sometimes even decreased it after success because he was afraid that the performance was "too good."

Differentiation of interests and values often arises on such a basis. The individual is chiefly interested in those fields in which he has demonstrated ability and is indifferent to achievements[22] in other fields. Feelings of failure are not experienced in the latter since he is not interested in them and moreover, he may generously admit his clumsiness in those fields, frequently implying at the same time, however, that such ineptness is a function of his lack of interest. E.g., *M*50 knows he's

> not so good in certain things but I don't feel failure. I can't play bridge very well and my dancing is pretty rotten. At least that's what they tell me. But I don't feel it's a failure. . . . I feel that I could do better and that I don't care enough about them to improve so much in them anyway. I'd just rather sit back and watch others play bridge. . . .

Thus the effect becomes the cause, or in other words, there is a "willing of the obligatory." The theoretical fields in which one might experience failure are thus narrowed down more or less to those activities where there are greater than chance occurrences of success. *M*21 exemplifies this type of protective reaction in its extreme, i.e., not only narrowing his field of interest and endeavors to a large extent but also tending to become satisfied with whatever he actually achieves. He only experiences failure when "I fail in something I *know* I can do" which feeling in no way threatens self-evaluation since this type of acknowledgment of "low" performance tends to actually heighten the successfulness of the past performances.

---

[22] Differentiation of interests and consequent different behaviors toward tasks differing in interest-values is another basis for low average difference intercorrelations.

The most extreme form of such a system of differentiation of interests is that found in individuals whose first reaction is to belittle anything new and untried. The eternal scoffer is often seen to completely reverse his decision, however, when he is made to engage in the task and discovers aptitude for it. *M*21 began playing the target with marked languor and by some freak managed to get a fairly high score about the sixth trial. He immediately "perked up" and remarked, "You know this is interesting." This is not to say that everyone's interests and values are based solely on ability. Cultural values, parental values, and given group values may override interests based on self-ability. Individuals differ, moreover, in the degree to which they adjust themselves to their own capabilities and limitations. The disparity between one's interests and one's abilities may sometimes be a clue to insecurity and maladjustment. *M*32, a very outstanding student, has developed values quite apart from his present abilities:

> . . . the things I excel in I regard as essentially unimportant and place much more importance on those things I can't excel in. . . . At Columbia they excel me in the economic field and I tend to make that paramount. They have certain advantages which I don't have and counterbalance all I have over them scholastically. . . . My most important values are for economic success—materialistic you might call it, based entirely on what I haven't got but want above all—to be able to do what I want, when I want.

Actually, of course, scholastic work is not unimportant to him because he does obtain feelings of success from excelling in school. This is however, an example of a hierarchical organization of values where, since economic want and insecurity characterized the first 15 years of his life, the supreme value is assigned to the major source of his anxiety feelings. He has, furthermore, cultural sanction for such assignment of values.

Another type of "going out of the field" has been discussed, in another connection, by Horney. She points out that anxiety "when based on a feeling of being menaced, easily provokes a reactive hostility in defense" (8, p. 74). In this regard it does not differ in any way from fear, which may equally provoke aggression. Some of the subjects evidenced their anxiety by turning against *E*, attempting to pull *E* "down," feeling sorry that *E* had to do all that writing of their statements. *L9* demonstrated this type of reaction to insecurity feelings to a marked degree. In the sixth trial of the steadiness test, which he had begun by informing *E* that he had not slept the night before and consequently was not responsible for what he did and that furthermore it was not a matter of ability, he suddenly discovered that his scores were getting better. He was very much pleased. "This is more fun really," he exclaimed. But this was a damaging admission, following as it did the increase in performance and he quickly added, "because I'm learning more about psychology than you are. I'm closer to the experimental situation—don't I have more control over the experimental situation than you?"[23] *M71*'s feelings of insecurity, for a number of reasons were not aroused fully until towards the end of the interview, or perhaps the effect was cumulative, only becoming noticeable to *E* and to himself when a certain pitch of intensity had been reached. When they were aroused, however, he frantically turned on *E*, belittling the experiment, scornfully demanding to know whether *E* really believed the boys had told her the truth, and that failing to get the reaction from E that would have assuaged his anxiety feelings, he attacked *E* more directly, pointing out that his sister "whom I'm sure is younger than you," had accomplished so much more, clearly implying "so much more than you ever could." "Are you content," he asked with fine disdain, "to spend the rest of your life in this miserable bare room? . . . Have you no ambitions? . . . Didn't your family bring you up any better? . . ." A few of the subjects expressed hostility toward the very things they desired. *H28* intermittently is keenly aware of his lack of scholastic abilities, often feeling that even if he had worked harder he wouldn't have done "very well" anyway. He attempts, nevertheless, to disparage ability.

---

[23] It is extremely interesting that he returned several months later to tell *E* triumphantly that he had been accepted to medical school—a rare honor for a third-year student.

> I know I'm not good in a lot of things but I don't worry about it. In fact sometimes I'm secretly proud of it. I think it would be dull to be good in everything.

This inversion of his inner concern is clearly illustrated by the inordinate pride he takes in one of the few accomplishments he has been able to discover in himself—"good taste." "I've been told I'm very conceited about my good taste but I don't care." Bravado and aggressiveness as reactions against his insecurity and anxiety are illustrated throughout the interview with him.

6. *Rationalization.* Rationalizations of events and motives are met with very frequently, the most common variety being the assignment of blame for "low" scores to sundry external factors such as the slip of the hand, the noise of a passing truck, or a hurdy-gurdy which unfortunately did not pass as quickly. Occasionally when an individual failed to respond to a "high" score with an increased estimate he would attribute the increase in performance to an easier set of questions. Failure was turned into success, in a few instances, by attributing the failure to the interference of another and superior accomplishment. The examples of the usual type of rationalization are too numerous and familiar to bear elaboration. *L*60, however, protected himself from failure by such a tightly-knit rationalization system to warrant extended description. In brief, the system consisted of ascribing "unexpected" and "unpredicted" failures simply to the fact that he had had an "off-day," or "just wasn't on" and of predicting future events in a way calculated to cover nearly all possible outcomes. Thus if someone asks him before a ping-pong game whether he will win or not, he answers, "If I'm on, I will, if I'm off I won't." This answer bears a striking resemblance to that old favorite, "Heads I win, tails you lose." Naturally he resented having to estimate in this situation since it put him in a position where he had to commit himself more definitely than he usually does. The character of his estimation was determined both by his previous behavior, where he had always attempted to avoid being taken unawares by any outcome, since such unawareness left him open to feelings of failure, and by this attitude: "When I give an estimate and don't make as much I get mad at myself, but when I underestimate I don't care." Thus it was less precarious for him to estimate the minimum he might expect to do the next time, or even lower than that, which he admitted to doing on occasions

"to make sure I would beat it." When he does recognize failure the feeling is less intense than a feeling of success because "You feel you can always do better the next time." Throughout all the examples of the way in which he copes with all undesirable facts, including the famous ostrich-technique, runs the suggestion of his "infallibility" in predicting the outcome of his future endeavors. If he should be unsuccessful in predicting, however, he invariably displays remarkable ability for turning an apparent defeat into a success, though the success be no more than that he had predicted failure. His ability for rationalization is illustrated with artistry by himself.

> I'm a fisherman first, last, and always. If I have a fish and lose him, why I was smart enough to hook him; if I get it in why I was very smart; and if I didn't hook him at all, then he wasn't there, or no one could have hooked him.

He also had recourse to another classic response, "I wouldn't be surprised if I was, but I don't think I am." Rather than give further illustrations, it would be more valuable to point out that although this situation and many others elicit "minimum" estimates from him, when he felt called upon to defend himself, specifically his faith in his ability to be a doctor, in the face of a poor scholastic record which might prevent him from entering any good medical school, his "estimates" were just as exaggerated in the opposite direction.

> . . . I'm convinced even if I go to the worst med school I'll come out a better doctor than some who went to the best. Or I might say better than all who went to the best school. . . .

One other illustration might be given. *L*48 attempted to rationalize his behavior and to deny from consciousness inner involvement with his performance scores. What he did was

> take the mean of the first few trials and then give a little less and make more and beat the clock. . . . Caution. I guess I thought I'd get a better score that way. I didn't want to see failure. I didn't like staring grim reality in the face.

Yet earlier in the interview he flatly denied any experiences of failure in this situation. In life admits to the same "cautiousness."

> Yes, I think I always put estimate lower. I don't know why, except if I have to estimate myself, and do better—then people

> get a better opinion of you. I never stop to think why these things work.

He attempted to avoid failure in the target by ostentatiously not trying to get any particular score. This reaction he dignified by the term "experimenting," claiming he often did it in bowling too, probably also when he is afraid he won't do well. His response, when told at the end of the experiment that his target scores were below average, corroborates this analysis of his behavior. He thought it was "funny . . . To think that I could beat everyone in making the lowest score. It's not everyone who could make the lowest." I.e., actually his estimates were reactions designed to prevent him from experiencing failure in his own and *E*'s eyes, as well as to give him some feelings of success because he could usually "beat" them. He also illustrates the lack of consistency in attitude, which is often the clue to the fact that the individual is rationalizing. I.e., the explanation for an event unpleasant to face is completely ignored when that event has a pleasant outcome. When he was asked what his reaction would be if his scores were average or below average, he vigorously denied that he would have any reaction since the scores meant absolutely nothing to him. In reply to what his reaction would be if his scores were above average, he said, "It would make me feel better . . . because I did better than anybody else." When *E* pointed out that such an answer was indicative of his concern about his standing, he was a little taken aback—"Oh you should be a prosecuting attorney. . . . You pin me down every time." *E: Then you really do care?* "You see you're embarrassing me now. I feel myself blushing and that isn't quite fair." In life he also resolves failure by "going out of the field" and by repression. "I try to get them [i.e., failures] out of my mind as quickly as possible and think of successes."

7. *Summary.* When all evidence to show that the average difference score is a device used by the individual to protect himself from failure has been reviewed, it is still possible that high estimates may be maintained because the individual is very confident he will do better next time (though such confidence on a conscious level is sometimes symptomatic of repressed insecurity). Moreover, estimates yielding zero difference scores may be maintained because the individual is not interested in his achievement, and minus scores may be maintained because the individual is afraid he won't

do well next time. For various reasons already discussed, minus scores are infrequent in this group, but when found they can in no way be related to a low level of self-evaluation, a genuine lack of interest, or a genuine low level of strivings. They may be differentiated from high difference scores in that they might represent a giving-into instead of a protest against the fear of failure. Unfortunately analysis of the subjects could not be made very deeply. There is, however, the suggestion that those in the lower extreme were in general more "cautious" in their behavior than those in the upper extreme. There is the further suggestion that those in the *L* group tend more to "take the easiest way out" of situations, both environmental and psychic, in which they find themselves. Thus *L*9 whose insecurity feelings could only be allayed by scholastic superiority (who, incidentally, could not understand why he always experienced such intense satisfaction from "*A*"—more intense than many he knew who didn't get *A*'s as frequently as he did), preferred to be in a group which gave him less competition because then his superiority would shine out more by comparison. We might add that the knowledge of his definite, unchallenged superiority would also have allayed his anxiety feelings and thus eased the intensity of his insecurity.

VI. DISCUSSION OF RESULTS

Frank (1), adapting Lewin's concept of "needs," advances the hypothesis that the average difference score (the difference between the estimates of future performance and the level of past performance in the same task), depends primarily on the relative strength of three needs: "(*a*) The need to keep the level of aspiration as high as possible regardless of the level of performance. . . .[24] (*b*) The need to make the level of aspiration approximate the level *of future performance as closely as possible . . . (c) the need to* avoid failure where failure is defined as a level of performance below the level of aspiration, regardless of the absolute goodness" (1, pp. 119-120). In a later paper, Frank (2) realizes that the need to keep in touch with reality is always present in every normal individual, that the very definition of normality implies contact with reality. This need cannot explain difference scores which deviate from zero, but only why such deviations are not greater than they are. According to Frank, the average difference score as determined by these three needs would "vary in a consistent manner among individuals regardless of the situation" (1, p. 120).

Our statistical results did not reveal very great generality of the average difference score, but it has already been pointed out that the absence of generality at one level of description is not proof of self-inconsistency or specificity on another level of description. In the light of the preceding discussion, however, one is led to doubt whether the postulation of the above "needs" (even to describe, to say nothing of explaining the fact that some individuals have greater difference scores than others) contributes anything.

Aside from the fact that the level of momentary strivings is not, on the whole, accurately mirrored by the estimates and consequently not by the difference scores, we have seen how the need to avoid failure, the most basic need elicited in this aspiration level situation, can be resolved by high, zero, or minus difference scores. We have seen that high difference scores do not necessarily indicate either a high level of self-evaluation, a high level of aspiration, or a high level of expectations, and that low or minus difference scores do not necessarily indicate a low level of self-evaluation, strivings, or expectations. The question naturally arises, if the estimate does not necessarily represent the expectation, the goal, the hope of the individual, what *does* it represent? This experiment has revealed that cultural mores, hopes, desires, fears, interests, and expectations influence the nature of the estimate, but for any given individual these factors operate differentially, i.e., these factors weight the estimate in different proportions for different individuals, and as a corollary, for the same individuals under different conditions. Until we can explore the possibility that under certain conditions certain types of individuals will respond more to one factor or group of factors than to another, the question cannot be answered more concretely. If one wished, one could arbitrarily call the verbal estimate the "real goal" as Hoppe and Lewin do, but it is difficult for the writer to see what is gained thereby. One must still define "real goal" and in so doing must point out that various factors weight the estimate differentially for different individuals.

---

[24] This need is understood by Hoppe as representing the need to achieve success.

On one descriptive level we could say that the high difference score and the low difference score group are distinguished by greater and lesser "cautiousness." Though on the whole this is startlingly true, such a distinction has little more than descriptive value since the explanation for differences in cautiousness is the real problem in distinguishing the two groups. The reasons for such personality differences *do not lie* in the differentiated strength of the two needs, as we are led to believe by Frank, since the relative strength of these needs may be the same in individuals differing markedly according to the quantitative results. Rather, the basis of such differences in overt behavior lies in the genetic development of the individual, the crux of the whole problem of "individual differences." In the light of these data, it appears to the writer an over-simplification and implicit denial of the problems involved, to describe individuals at opposite ends of the quantitative continuum of difference scores in terms of differential strength of the "need to avoid failure."

One may not attribute the "low" average difference scores to a greater need to avoid failure, as Frank does, since it seems fairly clear from the interview data that individuals in all three groups could be matched for apparent intensity of the need to avoid failure. The "need to avoid failure" seems to be no more inhibited in the *H* group as a whole than is the "need to keep the aspiration level as high as possible" in the *L* group, if by aspiration level we refer to the actual level momentary strivings. The difference between the two groups does lie in the methods used in order to cope with, avoid, and resolve this need to avoid failure.[25]

It is difficult to affirm a dichotomy between the need to succeed and the need to avoid failure, in any culture where achievement is phrased in terms of individual ability to succeed in every venture, and where there are internal or external sanctions against failing in anything, where a feeling of failure brings with it a deep feeling of shame, sometimes accompanied by violent self-repudiation [cf., an account of the Ojibwa reported in (12)]. Since an achievement which can be characterized as having avoided failure, cannot also, except under very special conditions, be described in terms of success, however, a psychological distinction between the two must be made. It seems fairly clear too, that the attitude towards future events must be conditioned by the relative emphasis placed upon one or the other need. It does not follow from this, however, that in order consciously to stress achieving success, either the need to avoid failure or the intensity of the reaction to failure must thereby be lessened. It is clear that the individual must adopt some compromise attitudes or mechanisms whereby he can steer clear of the shoals of failure and at the same time satisfy the cultural demands of achieving success. It is here that different definitions of failure and success play a major rôle. If the need to avoid failure is paramount, the individual may at the same time concentrate on achieving success, and if the need to achieve success is paramount, the individual can still have an intense fear of failure. It is possible that an intense fear of failure can goad on an individual to achieve all the available marks of success, in which case one has an inversion, on the conscious level, of the underlying motivation. Feelings of failure are not necessarily repressed or inhibited under such circumstances.

A dichotomy between these two needs must rest on the assumption that the two are independently variable. The need to avoid failure always operates, though it be repressed, and it may operate to keep explicit estimates high, or to keep them low. Which way it operates may depend on the personality structure of the individual, on such characteristics as cautiousness, willingness to gamble, flexibility, rigidity, etc., and possibly also on differential thresholds of resistance to disappointment. Not that the two extreme groups *qua* groups can be distinguished on the basis of greater and lesser responsiveness to decreases in performance or "low" performance scores, though individuals within each group can be. On an objective basis they can only be distinguished by the ability to tolerate wider disparities between performance and estimate. The quantitative difference between the two extremes would imply that the "high" group was for various reasons more resistant to decreasing their estimates and the "low" group for various reasons more resistant to increasing their estimates above their performance. Frank's assumption that performance below the estimate yields feelings of failure ignores the potency of rationalization whereby the discrepancy between performance and estimate may actually be utilized consciously as an incentive to increase performance. If Frank's two needs are to be retained at all, the apparent dichotomy must be understood as not being an absolute one; the relationship between the two needs is a "dialectical" one, i.e., one capable of revealing a "reconciliation of opposites." The individual who strives for success can in so doing inhibit partially the fear of failure, or an uninhibited fear of failure can act as a sufficiently strong irritant to produce aggressive attempts towards "positive" achievement. The individual who is consciously or unconsciously working towards avoiding failure need not necessarily have a more intense fear of failure, but in some cases this fear may lead to the adoption of purely defensive measures. The concepts of success, moreover, need not differ in these two individuals.

---

[25] It should be noted that different methods of coping with failure may result in differential effort being applied to the task at hand. Certainly if one feels that he is well protected from the feeling of failure, come what may, it is reasonable that he may have less of a drive to achieve a particular goal.

Frank's two "needs," which appear to be in the nature of *ad hoc* deductions evolved in order to explain the obtained quantitative results, are not only seen to be inadequate as explanations but actually hinder adequate interpretations of the data. As descriptions of the quantitative results they are only partially adequate. But after labelling "high" difference scores as "need to keep aspiration level as high as possible" and "low" difference scores as "need to avoid failure," are we any closer to an understanding of the difference between the two quantitative extremes—or of the function of "high" and "low" levels of explicit estimates?

Hoppe interprets changes in aspiration level, defined by him as the actual level of momentary strivings, since in his study no overt estimates were asked for,[26] as due to the "tendency to keep ego-level[27] as high as possible" (7, p. 36). Such a tendency, he claims, is realized by the achievement of success "through the highest possible aspiration level" (*ibid.*). At this point of our knowledge it is at least premaure, if not gratuitous, to assume an ego-level as the source of the obtained end-processes of motivation, and also to assume that we are measuring ego-level by the level of aspiration. There is no reason to assert that there must be a one-to-one relation between the tendency to keep ego-level as high as possible and the realization of the tendency in the keeping of aspiration level "as high as possible." This would be reasoning by analogy. The interview data presented in the above sections, particularly those parts indicating the reasons for maintaining a particular type of estimate and the obvious lack of relation between level of self-evaluation and level of estimates, or even between implied level of strivings and self-evaluation, indicate once again the dangers of such reasoning. Beyond this, to say that there is ego-involvement in this situation, howsoever we define the ego, is really saying very little since there are few situations where the ego is not involved. Aspiration level situations are differentiated from situations in general only in that ordinary pressures upon the individual are intensified. This means that there is provoked or elicited the individual's repertoire of defense mechanisms, as well as a variety of other attitudes toward the self and towards basic cultural factors.[28]

---

[26] Since his criteria of height of aspiration level were based on overt manifestations, however, and since we have found how misleading inferences from overt behaviors may be, the data he obtained may have been the prototype of the explicit estimates in this situation.

[27] Ego-level is defined by Hoppe as "the wide embracing goals of the person . . . which extend far beyond the single task . . . and are related to the self-regard of the '*S*'" (2, p. 287).

[28] The predominance of "high" difference scores, not the presence of them, seems to be most satisfactorily explained, at least at present, as a function of cultural stresses and values. In Chapter III we discussed the operation of such factors as cultural optimism, the obligatory nature of this optimism, the institutionalized technique of avoiding the present unpleasantness and failure by "weighting of the future" and the cultural differentiation of attitudes towards tasks labelled "motor" and "intellectual." There is one other factor which contributed to making the target difference scores the highest of all the tasks, i.e., the fact that the target was almost universally regarded as a "game," a situation where the boundaries of "reality" were less pressing and wishes could be more easily substituted for "facts." This means that the individual could more easily estimate closer to the natural

Many personality characteristics are strikingly revealed in this experimental situation since it allows for the expression of attitudes toward success and failure, toward competition, toward authority and regulations, and finally toward one's ability in reference to the hypothetical "average" person. This means that anxiety and insecurity feelings and one's reactions to these will also be evidenced. For example, if the individual is interested in marks *qua* marks because he regards them as symbols of his superiority or inferiority he will probably show great concern with the scores obtained in these tasks, and all other things being equal, will work hard in all of them. In regard to working hard, however, it is significant that practically all the subjects expended great effort, and that few were able to take this situation casually. Those who are intensely interested in gaining social approval, attempted primarily to gain the experimenter's approval. In one case this led to a rigid attempt to "hit it on the head," in another to demonstrate high aims and show his "perseverence." If the individual does not ordinarily experience intense feelings of failure or success, he does not tend to experience either or both in the laboratory. If his insecurity is expressed by being "aggressive" or being "cautious" and "reserved" in life, these attitudes are brought to light. Finally, such reactions as "relaxation after success," tendency to ride on one's laurels, "tightening up" and decreased efficiency after disappointment, ready boredom with repetition, acceptance or rebellion against authority, are also revealed.

The problem of analysis of the basis or source of surface-behavior is complicated by the fact that individuals differ in degrees of awareness, of inhibition and of repression, so that statements purportedly revealing inner feelings and attitudes have differential significance for the analyst. In brief, it is clear that the quantitative scores would not, on the whole, reveal genuine difference in aspiration level if the latter were defined not as the explicit estimates but as the actual level of momentary strivings motivating the individual.

---

maximum of the task and closser to his actual level of strivings. To put it in another way—since "high" estimates have been found at times to function as a substitute satisfaction for "low' performance scores, and since this culture gives approbation to those who *tried* to achieve high goals though they did not succeed in doing so, "high" estimates as a protective response could more easily function in a situation regarded as a "game," of not much import. On the other hand, of course, the absence of as much pressure on the individual in a game situation may decrease the need for protecting one's actual level of strivings from objective defeat.

VII. SUMMARY

1. The subjects of this study were 82 white male Columbia undergraduates, fairly homogeneous with respect to age and educational status.

2. Six tasks were selected on the basis of four criteria mentioned in the text. The tasks were: synonyms, addition, steadiness, symbol-digit substitution, cancellation, and target. These tests were administered to the subjects in two one-hour periods in accordance with a modified rotation scheme. Each subject was also given a two-hour interview at the end of the second session.

3. The quantitative measure of aspiration level was the average difference score, the difference between the average performance and average estimate in the same task. For each subject, with the exception of 12 subjects for whom there were only five average difference scores, there were six average difference scores. For each individual the median difference score and the average deviation (a measure of individual variability in difference score) from this median difference score was computed, all of the difference scores having been transmuted into standard scores for the purposes of our computations.

4. The corrected reliabilities of the performance scores on the six tasks range from .89 to .99; the corrected reliabilities for the corresponding difference scores range from .95 to .99.

5. Although preliminary investigation indicated approximately zero intercorrelations among performances on the six tasks, final results show that four of the fifteen intercorrelations are higher than .20 and two are —.36 and —.60. Internal evidence suggests that performance intercorrelations significantly deviating from zero influence the size of the intercorrelations among the difference scores.

6. A trend seems to be suggested for intercorrelations in difference scores to be higher for tasks which occur in the same session.

7. All but two of the intercorrelations between the difference scores are significantly greater than zero, the median intercorrelation being .29.

8. The correlation between individual variability in average difference score and median difference score is only .33, indicating that if subjects differ with respect to generality of aspiration level, as measured by difference scores, this difference is not primarily a function of height of level of aspiration.

9. A very strong tendency is indicated for the difference score to be positive, i.e., for the estimates to be higher than the performance. An attempt is made to account for this fact in terms of cultural factors.

10. There is no one-to-one relation between confidence and size of the average difference scores.

11. Although all of the subjects were individually tested and each subject remained ignorant of the performance level of the others (i.e., the situation was not in itself competitive), a strong competitive spirit (desire to be on top, get the highest score of all, etc.) was manifested by a great many of the subjects. This again points to the significance of cultural factors.

12. Neither success nor failure can be generally defined in terms of the absolute or relative achievement scores. Success and failure are only relative experiences to some subjects; others have absolute standards for success and relative standards for failure; and still others have relative standards for success and absolute for failure. Perfection, moreover, may be a sufficient, but is not a necessary condition of feeling of success. The individual definitions of success and failure explain the fact that individuals with low difference scores are not necessarly dominated by fear of failure, for to such subjects success may be differently defined than for individuals with high difference scores.

13. Feelings of failure are reported most frequently in connection with the three tasks having known proficiency limits—the synonyms, steadiness, and target tests. Feelings of success are reported whenever a zero score is approached in the steadiness test or a bull's eye attained in the target test. No feelings of success are reported in connection with the synonym test, but no subject even approached a perfect score on this test, the maximum score obtained being 18 out of a possible 33.

14. The disparity between estimate and performance is not always reacted to in terms of success and failure. Disparities are often reacted to as incentives rather than as demonstrations of failure. This makes for one of the few clear differences between the high and low group, the subjects of the high group seeking to make use of the discrepancy as an incentive, whereas the subjects of the low

group cannot permit themselves more than the smallest discrepancies because such disparity between estimate and performance bears to them the connotation of failure.

15. Reactions when the performance was below the preceding estimate varied from feelings of failure, disappointment, discouragement, and "weighting of the future," to various combinations of these.

16. Subjects react to feelings of failure (*a*) by trying to do better; (*b*) becoming discouraged, disorganized, losing the desire to go on; (*c*) overcoming them by thinking of success, making excuses, "weighting of the future"; or (*d*) by various combinations of these such as becoming discouraged but trying to do better anyway, making excuses but at the same time losing the desire to go on, etc.

17. The subjects were divided into three groups according to the size of the median difference scores of the individuals, giving a high (*H*), medium (*M*) and low (*L*) group. A study of the interview material available for these subjects was made in an attempt to throw light on motives involved.

18. An examination of interpretations of the experimental question *"What will you do next time?"* suggests that the clue to the difference which made for the quantitative differentiation of the three groups is probably not to be found in differences of interpretation of the question.

19. An examination of the meanings of the estimates to the subjects indicates that *quantitative differences in the size of the difference scores cannot yield an accurate picture of possible differences in true aspiration level, since the actual level of momentary strivings may be of the same magnitude in individuals widely divergent in difference scores.* There is also no evidence to indicate that even a knowledge of the inner conscious level of momentary strivings, let alone the explicit estimates, would reveal the level of goal-desires which actually motivates the individual. Among the various observable factors determining the nature of the estimates are: anxiety and insecurity feelings; desire to excel, to succeed, to avoid failure; actual level of momentary strivings and the disparity between this level and the explicit estimates; and general past experiences which

have helped determine the subject's personality organization and thus his reactions to such demands of inner and outer forces. Such evidence as was obtained concerning the goal strivings of the subjects does not indicate any differences among the three groups.

20. This situation chiefly elicits not the level of momentary strivings, nor the level of expectations though these enter in differentially, but the resources of the individual to establish himself and avoid failure. The particular techniques utilized by the subject are those he usually utilizes in any life-situation which offers his ego-integrity a similar challenge. In that sense the average difference score represents not the height of the level of momentary strivings, but a kind of protective response consistent with the individual's personality structure and past experiences. Protective mechanisms utilized by the subjects may be classified under (*a*) emphasizing the *relation* between performance and estimate, (*b*) emphasizing performance, (*c*) emphasizing estimates when the latter are high, (*d*) emphasizing estimates when the latter are low, (*e*) withdrawing from the field, (*f*) rationalizing.

21. Finally, we can offer no evidence one way or another to prove that an ascertainment of the "true aspiration level," which refers to the "momentary" not the "ultimate" goal-strivings, would in itself be a significant clue to the personality organization and "deeper strivings" of the individual. Not only is the relationship between ego-level and aspiration level an assumption to be tested experimentally at the present time, but the very concept of ego-level itself is still too vague and tenuous to allow for discussion in connection with the present experimental problem. There is no doubt that the aspiration level technique does bring to the surface certain significant personality characteristics. Whether it does so more easily and more completely than other laboratory methods is an experimental problem.

In view of our critical comments on existing methods in this field it is incumbent on us to suggest directions in which future research might move in order to overcome the present difficulties.

Two separate emphases can mark future research on aspiration level, i.e., studies on changes occurring in the range of estimates brought about by uniform external factors such as incentives, drugs, and so on, and studies of the changes in estimates occurring under systematically varied conditions *in relation to the individual's particular organization of motives, needs, frustrations,* and so on. Particularly in view of our present inadequate understanding of exactly what the estimates of future performance reveal, the first type of study will tend to contribute little or nothing. Such studies are likely to yield the same inconclusive results as most of the studies on "praise and reproof," since the crucial fact of the meaning of the task and of the praise to the individual's pattern of living is disregarded. We have found that some individuals respond to praise while others increase their efforts under conditions of reproof, but we do not know *why,* which after all is the important question for psychologists in particular and educators in general. Thus too, with the first approach we may find that estimates increase in some or all of the subjects under the influence of tobacco smoke, but we will be no nearer the understanding of our basic problem, which is the reason why the estimates change in one way for *Y* and in another way for *Z*. Our problem is not to study changes in estimates *per se*. The net effect of such a procedure is to divorce the estimates from their meaning, since their meaning resides in the organism which is ignored. Our problem is to study the subject in the light of the changes in estimates which are brought about not by experimentally introduced artifacts but by a genuine change in the relationship established between the organism and his immediate environment. We are then enabled to study the nature of behavioral changes brought about in general, in response to given affective changes in the experimental situation, and compare these with the responses manifested under other known conditions. The estimates are at once the inadvertent and not too reliable indicators of motives and attitudes and at the same time also provide a technique whereby the subject can reveal attitudes without necessarily being aware of them. That is, the very methods he may use to guard against revealing himself are in themselves revealing in this situation.

This experiment has disclosed the differential operations of certain psychological mechanisms in the act of verbal estimation of future performance. There is reason to believe that the same mechanisms would operate in non-verbal situations as well. In the type of relatively unstructured situation used here it is possible tentatively to classify the individual's type of protective reaction to a potentially catastrophic situation. How accurate such classifications are can be experimentally determined by means of a series of increasing ego-pressure situations. What changes occur in the behavior-reactions of the subject when the situations are made more acute, when e.g., failure is more imminent not alone from the *E*'s but from the subject's point of view and when success is something to be intensely striven for? What sequences of behavior are there in individual *Y* who evidences *X* type of protective reaction when he cannot utilize any of his usual rationalizations of failure? It must be remembered that changes in behavior-reactions and shifts in motivation patterns are revealed only in part by changes in verbal or other types of estimates of future performance. More important clues are obtained from the changes in the meaning of the estimates to the individual as well as in other overt behavioral manifestations. Determining what the fundamental characteristics of different types of reactions to danger and to the experienced catastrophe are, would no doubt bring us nearer the goal of prediction of individual behavior and allow us eventually to postulate the optimum conditions for the "good" development of intellectual and emotional integrity.

How stable, how deep-rooted is the characteristic of normality, the characteristic of *equilibrium* (which restricts the range of self-approbation and self-condemnation within bounds beyond which lie delusions of grandeur on one hand and suicidal attempts on the other), in individual *Z* who uses *Z* type of protective reactions? In other words, to what extent can *Z* be made to feel failure, to feel success, beyond the extent of his responsiveness to the original situation, before in one way or another he repudiates completely both *E* and the experiment?

REFERENCES

1. Frank, J. D. Individual differences in certain aspects of the level of aspiration. *Amer. J. Psychol.*, 1935, **47**, 119-128.
2. ———. Some psychological determinants of the level of aspiration. *Amer. J. Psychol.*, 1935, **47**, 285-293.
3. Garrett, H. E., & Schneck, M. R. Psychological Tests, Methods, and Results. New York: Harper, 1933.
4. Gould, R. Factors underlying expressed "level of aspiration." *J. of Psychol.*, 1938, **6**, 265-279.
5. Gould, R., & Kaplan, N. The relationship of "level of aspiration" to academic and personality factors. *J. Soc. Psychol.* (In press.)
6. Hausmann, M. F. A test to evaluate some personality traits. *J. Gen. Psychol.*, 1933, **9**, 179-189.
7. Hoppe, F. Erfolg und misserfolg. *Psychol. Forschung*, 1930, **14**, 1-63.
8. Horney, K. The Neurotic Personality of our Time. New York: Norton, 1937.
9. Jack, L. M. Behavior of the preschool child. *Univ. Iowa Stud., Stud. Child Wel.*, 1934, **9**, No. 33.
10. Lewin, K. Dynamic Theory of Personality. New York: McGraw-Hill, 1935.
11. Lynd, R. S., & Lynd, H. M. Middletown in Transition. New York: Harcourt, Brace, 1937.
12. Mead, M. Coöperation and Competition among Primitive Peoples. New York: McGraw-Hill, 1937. (Pp. 87-126.)
13. Murphy, G., Murphy, L. B., & Newcomb, T. M. Experimental Social Psychology. (Rev. ed.) New York: Harper, 1937.
14. Murphy, L. B. Social Behavior and Child Personality. New York: Columbia Univ. Press, 1937.
15. Murray, H. A. Techniques for a systematic investigation of fantasy *J. of Psychol.*, 1937, **3**, 115-143.
16. Page, M. L. The modification of ascendant behavior in preschool children. *Univ. Iowa Stud.: Stud. Child Wel.*, 1936, **12**, No. 3.
17. Thorndike, E. L. Teacher's Word Book of 20,000 Words. New York: Teach. Coll., Columbia Univ., 1932
18. Updegraff, R., & Keister, M. E. A study of children's reactions to failure and an experimental attempt to modify them. *Univ. Iowa Stud.: Stud. Child Wel.*, 1937, **13**, No. 4.

$7.00 per volume
Single numbers $3.00

QUARTERLY

May, 1939
Volume 21, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota

Charlotte Bühler
Universität, Wien

William H. Burnham
Clark University

Cyril Burt
University of London

Leonard Carmichael
Tufts College

Ed. Claparède
Université de Genève

Edmund S. Conklin
Indiana University

Sante De Sanctis
R. Università di Roma

Arnold Gesell
Yale University

William Healy
Judge Baker Foundation, Boston

Leta S. Hollingworth
Teachers College, Columbia University

Buford Johnson
The Johns Hopkins University

Harold E. Jones
University of California

Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal College

K. S. Lashley
Harvard University

A. R. Luria
Medico-biological Institute, Moscow

Toshio Nogami
Kyoto Imperial University

Henri Piéron
Université de Paris

George D. Stoddard
State University of Iowa

Calvin P. Stone
Stanford University

Lewis M. Terman
Stanford University

Godfrey Thomson
University of Edinburgh

E. L. Thorndike
Teachers College, Columbia University

C. J. Warden
Columbia University

John B. Watson
New York City

Helen Thompson Woolley
Teachers College, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

Provincetown, Massachusetts

Copyright, 1939, by The Journal Press

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the Act of March 3, 1879.

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879.

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the *request* of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out. Beginning with January, 1938, the following methods and terms will be in force:

(1). The size of each volume will be 400-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its *stated date*.

(5). Each author will be charged the actual manufacturing costs of any zinco or half tone cuts.

(6). The publication charges will be as follows:

4	page signature	....	$12
8	" "	....	16
12	" "	....	24
16	" "	....	32
20	" "	....	40
24	" "	....	48
28	" "	....	56
32	" "	....	64
36	" "	....	72
40	" "	....	80

These charges are *tentative* and will be adjusted as justified by the subscription list.

(7). The accumulated units will be mailed to subscribers quarterly. January, April, July, and October. The units may vary in size, to accommodate publication needs.

Beginning with January, 1938, eight hundred to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume QUARTERLY May, 1939
Single numbers $3.00 Volume 21, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

MAY, 1939

(Manuscripts are printed in the order of final acceptance)

SOME LIGHT ON THE PROBLEM OF BILINGUALISM AS FOUND FROM A STUDY OF THE PROGRESS IN MASTERY OF ENGLISH AMONG PRESCHOOL CHILDREN OF NON-AMERICAN ANCESTRY IN HAWAII 119
By Madorah E. Smith

Copyright 1939, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1939, 21, 119-284.

SOME LIGHT ON THE PROBLEM OF BILINGUALISM AS FOUND FROM A STUDY OF THE PROGRESS IN MASTERY OF ENGLISH AMONG PRE-SCHOOL CHILDREN OF NON-AMERICAN ANCESTRY IN HAWAII*

MADORAH E. SMITH

Teachers College, University of Hawaii

Acknowledgments 121

I. Introduction 123

II. Subjects used in this study 129

III. Methods 141

IV. Results 151

- A. Proportion of English used 151
- B. Errors in the use of English 167
- C. Length of sentences 187
- D. Analysis of sentences according to form and function 195
- E. Special analysis of questions 200
- F. Analysis of conversations according to parts of speech 206
- G. Sex differences in the mastery of English 213
- H. Home influence reflected in the children's speech 215
- I. Other environmental factors 232
- J. The effect of the presence of adults only 241
- K. Results from the repeated study of forty-four cases 244
- L. Is the handicap due only to Pidgin English or also to bilingualism 246
- M. Type of words in the mother tongue that linger longest when it comes in contact with another language 253
- N. Commonest English words 257

V. Summary and conclusions 265

Appendices 273

References 283

---

*Received in the Editorial Office on November 11, 1938, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

ACKNOWLEDGMENTS

As this study involved the collection of material in seven distinct languages, not counting dialects it would have been impossible without the aid of bilingual students of the University of Hawaii, either through the contribution of material gathered for their own masters' theses or through painstaking and interested coöperation in *F.E.R.A.* and *N.Y.A.* projects or otherwise.

The persons without whose aid this study could not have been completed, and whom I wish to thank for help in collecting, analyzing, and interpreting, are: *Chinese data*—Miss Elinor Chun, Mrs. Elsie Y. K. Motoyama, and Miss Alyce Yee; *Filipino data*—Messrs. Roman Cariaga and Anastacio Luis; *Hawaiian data*—Miss Alice Keakealani Lee and Mrs. Sadie Kaheaku Ragbir; *Japanese data*—Misses Elsie Hokada, Namiko Iwashita, Sakiko Okubo, Dorothy Sakamoto, and Mr. Jitsuichi Masuoka; *Portuguese data*—Miss Dorothy Vierra and Messrs. Enos Medeiros and Abel Pinto; *Korean data*—Misses Gladys Kim, Alice K. Lee, Ellen Yang, and Messrs. John Kwon and Arthur Song.

For aid in analyzing, typing, and statistical work—Misses Florence De Mello, M. Elizabeth Smith, Marion Wong, and Messrs. Nyuk Shin Ching, Frank Rapozo, and Sadaji Takasaki.

Thanks are also due to Drs. Willis Coale, Denzil Carr, John Reinecke, and Andrew Lind for advice and criticism of the manuscript.

MADORAH E. SMITH

TEACHERS COLLEGE
UNIVERSITY OF HAWAII
HONOLULU, HAWAII

I. INTRODUCTION

A. THE PROBLEM

The purpose of this investigation is threefold. First, it is a survey of the comparative progress made in the mastery of the English language on the part of young children of different racial antecedents in the Territory of Hawaii, so that we may know better what to expect of the children of different racial groups at the time of school entrance and the Americanization of each group, so far as that may be measured by the use of its national language.

Second, it is an attempt to study a number of factors that may hinder or further the children in such mastery.

Third, since the languages spoken in Hawaii are many, and many of the children are bilingual, it is an attempt to throw a little light on the problem of whether or not bilingualism is a hindrance in the mastery of speech to very young children.

Our subjects are, therefore, all children either below school age or in their first year of school. All are between 18 and 78 months of age.

The survey was made by securing samples of conversation of a thousand such children who might be considered representative of this section of the population of the Hawaiian Islands.

B. THE LANGUAGE SITUATION IN HAWAII

Owing partly to its position at the crossroads of the Pacific, but more to the need for laborers on its plantations, Hawaii has come to have a polyglot population.

Even before the annexation of the Islands to the United States, English was taught in nearly all of the schools (2, pp. 41-3). After that event, two official languages, Hawaiian and English, were recognized. Although since then, the use of English has steadily increased and that of Hawaiian has decreased, election ballots are still printed in both languages.

Since 1876, various groups of immigrants have been induced to come to Hawaii to labor on the plantations. The first of these were the Chinese (2, p. 13), who came between 1876 and 1897; next were the Portuguese, who came in two waves, the larger in 1878 to 1887 and the smaller from 1906 to 1913; then came the Japanese from 1885 to 1919; almost all the Puerto Ricans came about 1901, and the Koreans in 1904 and 1905; then came the Spanish from 1907 to 1913; and lastly came the Filipinos, who began arriving in 1907 and have only recently ceased coming. Each group came using a different language and was at first unable to converse with the other or with the Hawaiians or the English-speaking plantation owners. To meet the demand for a means of communication between immigrants and others, a local hybrid language known as "*pidgin*" has developed. Reinecke (26) considers this hybrid tongue to be really a continuum from a creole dialect, such as readily arises in a plantation environment to meet the needs of communicaton between a dominant and a servile class, which is spoken by the less literate foreign-born, to a colonial dialect, that is more like standard English and that is used by the majority of the non-"haoles"[1] less than forty years of age.

Children of Hawaiian ancestry and those whose ancestors came from each of the immigrant groups mentioned above, except the Puerto Ricans and Spanish, are included in the present study. No attempt was made to include these two groups—the Spanish, because it is now a very small one in Hawaii, numbering in 1930 only 1,219 out of a total populaton of 368,336 (2); the Puerto Rican, because although it was in 1930 a slightly larger group than the next smallest, the Korean, there was no Puerto Rican student at the University of Hawaii to aid in the research.

The languages spoken by our subjects, then, include English and its variety, "pidgin English"; Hawaiian; several Japanese dialects; two Chinese dialects, Hakka and Cantonese; three Filipino dialects, Tagalog, Ilocano, and Visayan; Korean; and Portuguese.

It is desirable to make a few comparisons among these many languages.

As Reinecke and Tokimasa (27) picture the colonial dialect, it is composed mainly of English words and about a thousand Hawaiian words and phrases for which there is no exact English equivalent, as is the case of names for much of the native flora and fauna and of words referring to things peculiar to Hawaiian life and culture, childish technical terms, and colorful words.

As to its grammar, there is practically no inflection of verbs, all modifications to indicate mood, time, or voice being made by a few auxiliaries such as *"been," "stay"* (which is from the Portuguese *"esta"*), *"no can,"* and *"try,"* the last-named as an auxiliary with the imperative. Nouns have no possessive, and there is considerable divergence from standard English in the use of the plural. Reinecke lists peculiarities in the use of all other parts of speech which might be summed up as tending in the direction of reducing what few inflections standard English has, omission of many words considered essential, changes in word order, particularly in the case of modifiers and in interrogative sentences, and in the overuse of some favorite words due to extension of their meaning, probably as a result of a small vocabulary.

---

[1] "Haole" is an Hawaiian word that has come to be used to apply to almost all Caucasians except the Portuguese.

Dialects of the other languages will not be discussed; but a brief description will be given of the methods by which the tense, voice and mood, person and number, are expressed in verbs; number, gender, and case, nouns and pronouns; the way modifiers are compared; the use of articles, connectives, and the copula; and the expression of negation and interrogation in each of the six languages.

Portuguese (9) is the most like English of the languages under discussion, although it is a more highly inflected one. Some words in both languages come from the same Latin roots. Verbs have three regular conjugations, being inflected by endings. The passive voice is rarely used, but it has reflexive and periphrastic voices that English does not. Two auxiliaries, *"estar"* and *"ser,"* are used in combined forms and sometimes are used like the English copula. Pronouns have the same persons, case, and genders as in English, except that there is no neuter pronoun. Nouns, adjectives, and articles all are inflected as to number and gender. There are both definite and indefinite articles, but the latter is the same as the adjective "one." Adjectives, though occasionally compared by endings, are more often, especially in spoken language, compared by separate words and at times colloquially by both.

The use of connectives is quite similar to that in English. There are several negative adverbs, but the word used where English would answer a question by "no" is also used where a different word, "not," would be required in English. Unless questions begin with interrogative words, they are distinguished from declarative sentences only by intonation. Questions seeking corroboration may be asked as in English by using a phrase meaning "is it not?"

The Filipino dialects (4) spoken in Hawaii are similar in grammatical structure, though they differ much in vocabulary. The Roman alphabet is used as in English, and they have borrowed from Spanish, especially in religious terms and some constructions, Sanskrit, and Chinese. In Tagalog (which will be used as a sample dialect), the verb has imperative, infinitive, and indicative moods, four tense forms, active and two kinds of passive voices. Auxiliaries and infixes are used to express these different forms.

There is no inflection to denote gender, person, and number of nouns or, with a few exceptions, case of pronouns. Personal pronouns are altered to express case and number but not gender. Plural is expressed by "*manga*" or "*mga*" before and by the form of the definite article. Articles are inflected, and the indefinite article is the word used as the numeral "one." Adjectives have the plural idea expressed by "*manga*" or reduplication and are compared by the use of different words. There are many negative adverbs, but there is none that answers a question in the negative that could not be used where English uses "*not.*" Besides the connectives used in English, "ligatures" or particles used to connect modifiers with the word modified occur. Questions may be asked either by intonation or by the use of interrogative adverbs. Negatives come first in negative sentences and verbs first in affirmative sentences.

Hawaiians (3) had no written language when the American missionaries came, so when the latter reduced Hawaiian to writing, the Roman alphabet was used. It has no inflections, but all the varieties of tenses, moods, etc., used in Hawaiian are formed by particles. These forms are less definite than in English, and the verb is impersonal. Pronouns have three persons and three numbers, singular, dual, and plural. Case is expressed by prepositions for both nouns and pronouns; gender either by different words or by the use of adjectives meaning male and female. The plural of nouns is expressed in several ways, by a plural article, plural signs, accent, or, after proper nouns, sometimes by the syllable "*ma.*" There are three articles, an indefinite and both singular and plural definite articles. The numeral "one" may be used instead as an indefinite article. Adjectives are compared by means of adverbs. There are no separate negatives for the English "no" and "not." Where English uses the copula, the arrangement of words or the affixing of a particle serves the purpose in Hawaiian. Interrogative words are used, besides which the negative phrase *"aole anei"* begins the sentence when the answer "no" is expected; otherwise, the leading word is followed by *"anei."* Emphasis is expressed by the position of words.

The other three languages all have their own peculiar method of writing. Korean and Japanese have alphabets, but both make some use of Chinese ideographs.

Meriggi (21) and Carr (7) have both pointed out that it is erroneous to consider spoken Chinese a monosyllabic language, and the *Encyclopedia Brittanica* (10) suggests that its use of couplets is so frequent that it might be considered disyllabic. There is no inflection unless the use of the same plural particle to express number of all three personal pronouns and of the same syllable to express the genitive case of each be so considered; otherwise, use is expressed by preposition or position. Although voice, tense, and mood can be expressed by particles, they are not much used, but such ideas are to be gathered from the context. In written language, no character can be used only as one part of speech, but in spoken language they are not so interchangeable.

Williams (35) classifies the parts of speech, as do Chinese grammarians, into essential words: namely, nouns and verbs, and empty words which include conjunctions, interjections, introductory words, finals, and particles. There is no difference between adjectives and adverbs, and there are no articles. A question cannot be answered by a single negative word, and the negative used in such an answer would also be used where English uses "not." Questions are asked by interrogative words and the addition of one of several syllables, corresponding to English *"eh?"*, at the end of the sentence. A peculiarity is the use of classifiers before nouns after numbers; and of enclitics, syllables used only to express a grammatical function.

Japanese (20) has different styles of the verb according to the person addressed. Verbs coming from the Chinese are conjugated by auxiliaries. In the polite form of the Japanese verbs there are seven tenses. The passive is used in quite a different way than in English. There is an imperative but no infinitive, and the verbs are impersonal. The subject is often omitted, especially where in English a pronoun would have been used. Nouns rarely have number or gender. Case is expressed by post-positions. There are no articles, no relative pronouns, and but few real adverbs. Adjectives are quasi-verbal and are conjugated like verbs. Interrogation is expressed by postpositions. The syllable *"ne"* at the end of the sentence has a use quite similar to *"n'est-ce pas"* in French. There are no words corresponding to the English "yes" or "no." *"Iie"* which means "no" is little used except in indignant denial. After a negative question the use of "yes" and "no" is exactly opposite to English usage. Interjections and honorifics are plentiful. Ellipses are frequent. Negation is expressed with the verb by a special conjunction.

Korean (*33*) has been influenced by both Chinese and Japanese. Like the latter, it has three styles according to whether an inferior, equal, or superior is addressed. The verb has three voices—active, passive, and causative; it has four simple and five compound tenses; it has different endings when expressing negative and interrogative ideas. There are imperative and volitive moods. There are participials and verbal nouns. Adjectives are quasi-verbs and are conjugated either positively or negatively. Special classifiers are used with numerals as in Chinese. Besides the interrogative conjugation, there are interrogative pronouns which may be used in asking questions. Adverbs are rare. The force of connectives is expressed by postpositions and relative participles. There are no articles.

This brief summary of languages spoken by our subjects may be of help in explaining their errors in the use of English.

II. SUBJECTS USED IN THIS STUDY

A. SOURCE

The children whose conversations were used in this study are representative of all the major racial groups found in the Hawaiian Islands. Every one of them is an American citizen, having been born in the United States; but for convenience, they will be referred to by the name of the race or nationality of their ancestry.

The majority of the children were from the city of Honolulu, but one group of 125 Filipino children was taken from plantation camps and villages in rural Oahu and Maui. The Honolulu children were selected at random from all parts of the city, but in such a manner that each major racial group consisted of 25 children at each age level from two to six years[2] and approximately the same number of boys and girls.

Seven such groups of 125 children were secured, representing the Chinese, Japanese, Korean, Hawaiian and part-Hawaiian combined, Portuguese, and Filipino population of Honolulu, and the one rural group of Filipinos on plantations. Besides these main groups, there were 100 children of varied racial origin, who were studied while at kindergarten instead of in their own homes as the rest were; and another mixed group of 25 children who, although observed in their own homes, were not studied in the standard situation—while at play with other children—but when alone with adults.

This total of 1,000 cases furnished the material for the main study. A few other miscellaneous records, contributed by persons who had taken them in the course of other studies, have been used for certain comparisons. These were of four rural Hawaiians studied by Miss A. Keakealani Lee, of seventeen rural Japanese observed by Mr. J. Masuoka, and one record of a seven-year-old included in the study of repeated records.

Comparisons will be made frequently to a group previously studied of Caucasian English-speaking children who knew but the one language (28, 29, 30, 31). This group, most of whom were observed in Iowa, will be referred to as Caucasians. A few of these children were of non-Portuguese ancestry living in Honolulu, to whom others were added from time to time after the previous study was com-

---

[2] Two years included all children from 18 months, 0 days, up to but not including 30 months, 0 days.

TABLE 1

PERCENTAGE OF PARENTS OF THE CHILDREN ENGAGED IN EACH OCCUPATIONAL CLASS AS COMPARED WITH PERCENTAGES OF MEN OF THE SAME RACE SO ENGAGED IN THE HAWAIIAN ISLANDS ACCORDING TO THE U. S. CENSUS OF 1930

	Number	Professional	Clerical	Public service	Trade	Transportation	Domestic service	Manufacturing	Extractive	Forestry & Fishing	Agriculture
Chinese											
Honolulu		3.9	15.6	1.9	29.9	4.7	18.9	18.9	0.3	0.4	5.4
Parents	120	6.7	18.3	2.5	31.7	1.7	18.3	20.0	0.0	0.0	0.8
Japanese											
Honolulu		4.0	4.6	0.6	16.0	10.0	17.8	34.8	0.3	5.1	8.9
Parents	125	4.0	0.6	0.8	17.2	4.0	18.4	40.0	0.0	2.0	3.0
Korean											
Honolulu		3.6	2.7	2.0	12.4	8.8	26.6	31.5	3.0	0.0	9.3
Parents	104	5.0	9.6	8.7	14.4	8.7	13.5	37.5	0.0	1.0	2.0
Hawaiian											
Honolulu		5.7	9.3	16.1	6.2	27.8	3.4	29.5	0.2	0.9	0.9
Parents	125	3.2	7.5	19.2	6.4	38.3	1.1	24.5	0.0	0.0	0.0
Portuguese											
Honolulu		2.5	8.5	8.5	12.3	20.1	4.8	40.1	0.2	0.4	2.8
Parents	100	3.0	5.0	6.0	11.0	31.0	1.0	35.0	2.0	0.0	6.0
Filipino											
Honolulu		3.2	1.3	5.6	8.9	13.2	21.5	36.4	0.2	1.0	8.9
Parents	96	2.1	0.0	14.7	9.5	17.9	22.1	28.4	0.0	2.1	2.1
Plantation Filipino											
All Hawaii		0.6	0.2	0.8	1.3	2.2	2.0	8.4	0.0	0.5	84.0
Parents	125	0.0	0.8	0.8	4.0	0.8	0.8	6.4	0.0	0.0	86.4

pleted. This Honolulu group, which is used a few times for comparison, will be referred to as "haoles," an Hawaiian term now frequently used to designate people of such antecedents.

Each of the seven main groups studied was composed almost entirely of children of pure racial ancestry, except for the Hawaiian, which consisted of children of all degrees of Hawaiian blood who were from one-fourth to entirely Hawaiian ancestry. There are also three exceptions in the Chinese group, where the fathers were Japanese or Caucasian—but the children heard only English and Chinese in their homes and had practically no contact with their non-Chinese relatives. The Caucasian father had left his half-Chinese offspring.

B. AGE AND SEX OF SUBJECTS

The mean age of each of the seven main groups is approximately 48 months, and the standard deviation between 16 and 17 months (see Table 2). The greatest difference in age between the groups is 1.1 months, between Hawaiian or Portuguese and the plantation Filipinos. This is not a significant difference.

An attempt was made to have nearly the same number of boys and girls in each group. The difference is greatest for the Chinese, where there are 13 more boys than girls. There are 23 more boys than girls in the 875 children included in the seven main groups. The boys average 1.3 months older than the girls, but for no race is the sex difference in age significant.

C. REPRESENTATIVENESS OF SAMPLES AS TO SOCIO-ECONOMIC STATUS, BIRTHPLACE, OCCUPATION, AND EDUCATION OF PARENTS

A comparison of the per cent of the parents of the children engaged in different occupations with the per cent of the respective race engaged in such occupations according to the United States census data for 1930 (34) shows fair correspondence, which indicates that our cases are reasonably representative of their respective racial groups so far as sampling of different occupations is concerned. These data are shown in Table 1.

The basis on which the percentage of parents in each occupation was calculated was the number of those for whom accurate information was available. As the material for certain groups was gathered

TABLE 2

COMPOSITION OF THE GROUPS STUDIED

	Hawaiian	Chinese	Portuguese	Japanese	Korean	Filipino Honolulu	Filipino Plantation	Av. all
Parents of children								
Per cent born in U. S.	96	75	94	41	63	18	2	56
No. of years of English schooling	7.1	6.1	5.9	2.7	4.9	4.1	3.2	4.9
Average scores on Barr Rating Scale	6.54	9.62	6.91	8.62	7.39	6.23	5.50	7.26
English Language rating	3.0	2.5	3.0	2.0	2.5	2.2	1.9	2.3
Favorable home environment rank	1	2	3	5	4	6	7	
Racial efficiency index (Porteus)	51	83	60	86		33†		
Socio-economic status (Masuoka)*	4.2	3.5	6.5	4.5	8.5	10.0†		
Children								
Average age in months	48.4	47.8	48.4	47.7	47.6	47.7	47.3	47.8
σ of distribution	16.7	16.9	16.8	16.1	16.7	16.5	16.6	
σ of mean	1.4	1.5	1.5	1.4	1.5	1.5	1.5	
								Total all
No. of boys	60	69	64	66	62	63	65	449
No. of girls	65	56	61	59	63	62	60	426
Number of children attending kindergarten or school	8	42	6	58	25	18	2	139
								Av. all
Av. no. of mins. required per record	41	51	55	44	45	45	37	45

TABLE 2 (*continued*)

	Hawaiian	Chinese	Portuguese	Japanese	Korean	Filipino		Av. all
						Honolulu	Plantation	
Average age								
of boys	49.7	48.7	49.1	48.3	48.9	48.1	46.2	48.4
of girls	47.1	46.9	47.6	46.8	46.4	46.3	48.0	47.1
σ dist. of age								
for boys	17.2	16.7	16.6	15.4	15.9	16.2	17.0	
for girls	16.0	17.1	17.0	16.7	17.8	16.8	16.0	
σ mean								
for boys	2.2	2.0	2.1	1.9	2.0	2.0	2.1	
for girls	2.0	2.3	2.2	2.2	2.3	2.2	2.1	
Mos. diff. in age	2.6	1.8	1.5	1.5	2.5	0.8	1.8	1.3
σ of diff.	3.0	3.0	3.0	3.0	3.0	3.0	3.0	
Diff. ÷ σ	0.9	0.6	0.5	0.5	0.8	0.3	0.6	
Per cent of race that speak English (1930 census)	96.7	77.5	94.8	70.1	70.4	46.0†		

†All Filipinos.

*Average of Chinese and Haole rated, Hawaiian and part Hawaiian ranks averaged.

during the depression, the only occupation named was in 56 instances *F. E. R. A.* Even excluding these, the higher percentage of parents engaged in transportation in several cases is probably due to the large number engaged as road laborers under the *F. E. R. A.* who did not specify that this was not their ordinary occupation. There were eight deceased parents for whom it was not possible to learn their former occupations. The remaining omissions are due to the fact that in several cases only a large firm that employs many types of labor was named as the source of employment, and the interviewer could not obtain sufficient details as to the nature of the work to classify the occupation.

In most groups, there are fewer parents engaged in agriculture than is true for their racial group in Honolulu. This is due to the fact that the area for census data includes quite an extensive rural district from which very few of our cases were drawn.

Our Korean group seems to be somewhat highly selected, but otherwise the differences are not excessive.

Where the information was sufficient, the parental occupations were rated on the Barr Scale (32, pp. 63-69), which was constructed in such a manner as to give an index of the intellectual level characteristic of the occupation. The average for all seven racial groups is 7.26 (see Table 2) or somewhat below that of 7.92, estimated by Barr as the average for all men in the United States. It is probably true that the occupational level of the entire population in Hawaii, when the Other-Caucasian group is excluded, would be lower than in other parts of the United States, for as Lind (17) has pointed out, the Other-Caucasians hold a superior position in the professions and administrative lines.

The relative standing on the Barr scale of the different groups of parents is Chinese, Japanese, Korean, Portuguese, Hawaiian, and Filipino (see Figure 1). As both Adams (2) and Lind (17) have pointed out, the earlier comers to the territory found superior opportunities more easily. Therefore, the Barr rating as an indication of intelligence is not a valid measure for comparison between racial groups but only within groups.

The five races to whom Porteus and Babcock (25, pp. 108-112) assigned "social efficiency indices" in their book, *Temperament and Race*, follow the same order according to this index as they do in this study on the Barr Scale, except that the Japanese and Chinese change places.

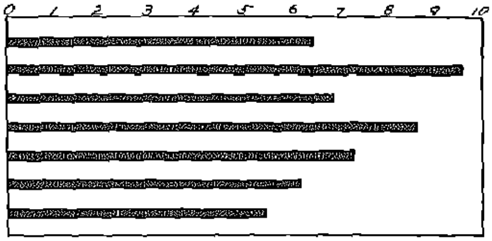


FIGURE 1

Average Barr Ratings of Occupations of Fathers of Children Studied

Column	Average rating of	
1.	Hawaiian	6.54
2.	Chinese	9.62
3.	Portuguese	6.90
4.	Japanese	8.62
5.	Korean	7.39
6.	City Filipino	6.23
7.	Rural Filipino	5.50

The racial groups follow the same order as to the percentage of parents born in the United States as they do as to percentage of instances of both parents native in Adams' (1) study of parentage of infants born in Hawaii in 1930, except that the Korean and Japanese change places. Our percentages are much higher than his,

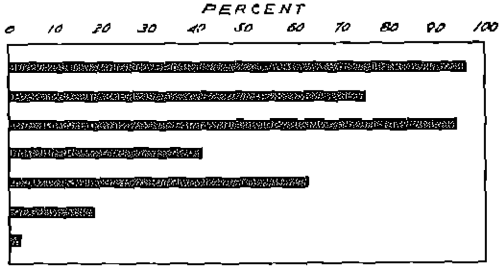


FIGURE 2

Per Cent of Parents Born in the United States

Bar		
1.	Hawaiian	96%
2.	Chinese	75
3.	Portuguese	94
4.	Japanese	41
5.	Korean	63
6.	City Filipino	18
7.	Rural Filipino	2

but they include those cases where but one parent was native-born. Moreover, his data are for all Hawaii, while ours for six of the seven groups are for Honolulu; and nearly all the foreign-born population of the Islands came as field laborers, so that Honolulu has a higher percentage of native-born than do the rural areas (Figure 2).

Masuoka (18, p. 51) had 50 Chinese and 42 Caucasians rate the racial groups represented in the Islands according to socio-economic status. The average of the ranks received from each of the two sets of raters given to the six racial groups in our present study is shown in Table 1. To get the Hawaiian rating, the average ranks received by the three sub-groups of Hawaiian and part-Hawaiian he studied (all of whom are represented in our Hawaiian group) were averaged together. Comparison with our ranks on the Barr rating scale show that our Hawaiian group is a little low, the Korean too high. The Hawaiian group studied seemed to be more affected by the depression than the others, and the occupational rating was less complete.

The average education of the parents of the different racial groups follows very much the order of arrival of the respective races in the Islands. The Japanese average less English education than do the Filipinos, who are later arrivals in Hawaii, but as the language of the public schools in the Philippines is English, that is an exception to be expected. The Koreans have had more English education than the Japanese, although they arrived in the Islands a little later. This is another indication that they are a selected group.

D. LANGUAGE OF HOME AND OTHER ENVIRONMENTAL FACTORS STUDIED

The language spoken in the home was rated for each child according to information supplied by the interviewers. If correct English only was used in the home, the rating was five; if good English and another language were spoken, four; if, besides a foreign language, both correct and pidgin English were heard by the child, as in the case where one parent only spoke correctly, the rating was three; if only pidgin English, two; if only a foreign language, one (Figure 3). The rating is, of course, crude and subject to error, as the interviewer did not often hear all members of the family speak. But there is a close correspondence in the ranks of racial-group averages by this rating and the ranks of similar groups of university students,

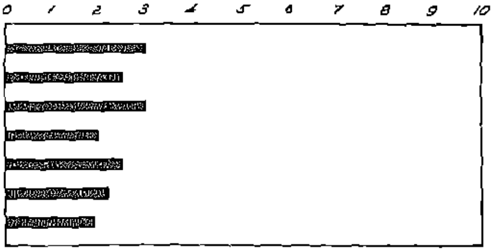


FIGURE 3

AVERAGE LANGUAGE RATINGS OF HOMES OF CHILDREN STUDIED

Bar		
1.	Hawaiian	3.0
2.	Chinese	2.5
3.	Portuguese	3.0
4.	Japanese	2.0
5.	Korean	2.5
6.	City Filipino	2.2
7.	Rural Filipino	1.9

according to averages found on the Hofmann *Bilingual Inventory* (Table 3). Correlations of +.58 to +.74 were found between

TABLE 3

LANGUAGE RATINGS OF RACIAL GROUPS COMPARED WITH SCORES OF UNIVERSITY STUDENTS ON THE HOFFMANN BILINGUAL INVENTORY

	Chinese	Japanese	Korean	Hawaiian	Portuguese	Filipino
Language ratings of children's homes	2.5	2.0	2.5	3.0	3.0	2.0
Scores on Hoffman Bilingual Inventory	12.6	21.4	17.0	5.3	†	

†The Portuguese were grouped with the other Caucasians in the University study. This group scored 1.4 on the Hoffman Inventory. There were no Filipinos rated in the University study.

English education of parents and language rating for those groups where the ratings differed enough to make the calculation of the correlation of any value. Almost all the Filipino homes were rated "2," that is, the child heard only pidgin English. The slightly lower average of plantation Filipinos is due to a few more homes where a Filipino dialect only was spoken. The Japanese average is the same as that for both Filipino groups taken together, but the variability was much larger. There were many homes where Japanese was preferred and a corresponding number where at least one member of the family spoke reasonably correct English.

Although the Koreans came much later to Hawaii than the Chinese, the average rating of their homes is the same. Since the average education of parents is lower, the Hofmann score of the university students is higher and the correlation between language and education lower, it would appear that the Korean interviewers were too lenient in recording the home language. Yet the study shows the proportion of English used by the young Koreans is greater than that used by the Chinese and the amount of error in its use not much greater. So the conclusion is that the Korean interviewers were only slightly more lenient than the Chinese (who were master's candidates) in marking as pidgin less often the English used in the home. These two groups both showed wide variability, but the child in the average home hears some correct but more pidgin English and some foreign language.

The typical Hawaiian or Portuguese child of the groups studied hears both correct and pidgin English. These groups may, therefore, be considered not as truly bilingual groups, for in very few homes was either Hawaiian or Portuguese used to any considerable extent, but rather as samples of children whose English is decidedly contaminated by pidgin so that it is possible to use these groups in estimating whether the retardation in speech noted is due to bilingualism or to the incorrect English current in Hawaii.

Adams (2, p. 43) gives the percentage of people in Hawaii over 10 years of age who, in 1930, could speak English as 96.7 (average of Hawaiian and part-Hawaiian) for Hawaiian, 94.8 for Portuguese, 77.5 for Chinese, 70.4 for Korean, 70.1 for Japanese, and 46.0 for Filipino. The groups over 90 per cent of whom spoke English are rated three in this study: that group about three-fourths of whom spoke English was rated 2.5; the two groups speaking least English received the lowest ratings. The Koreans are rated higher, 2.5, although in 1930 only less than one per cent more of them spoke English than did the Japanese. Two reasons for this are suggested: the Koreans are a select sample; they are a relatively small group, do not live in districts where most of the residents are of the same race as do the Japanese, and are not concerned about their children speaking the parental language (for much less often are they sent to foreign language school); therefore, if they can speak English at all, it is more often used.

In order to determine which group had the more advantageous home background for learning English, the sum of the ranks from most to least fortunate as to English language background on the four factors of parental birthplace, occupational level of father, language rating of home, and extent of English education was found. According to this method, the order for favorable home environment in learning to use English is as follows: Hawaiian, Chinese, Portuguese, Korean, Japanese, and Filipino. But if the groups are ranked also as to number in attendance at kindergarten or school (either at nursery school or just entered the first grade), the Chinese would change places with the Hawaiian and rise to first place.

In order to determine if the various factors considered in the study of the children's parents were related, and also as a check on the validity of language ratings of the home, correlations between Barr ratings of father's occupation, years of English education of mid-parent, and language rating of home were calculated. Language rating of home and English education of mid-parent gave the highest correlations, ranging from .58 to .74, or with parental occupation partialed out, from .55 to .68, the lowest relation being in the case of the Hawaiian group (see Table 4). As that group used so little but English in the home, the rating resulted in very coarse categories. Except for the Hawaiian and Portuguese, the correlations between occupation and education which range from .20 to .64 almost vanishes when language is partialed out, while for the groups using more English, the resulting correlations are .47 and .53.

For the Chinese and Japanese, facility in the use of the English language seems to be more closely related to a higher occupational status than does extent of English education; for the Korean, Portuguese, and Hawaiian, the relation between education and occupation is more important, probably because the groups as a whole are reasonably facile in the use of English.

Data were secured as to the attendance of the child at kindergarten or school, and his order of birth, so that these factors might be studied. Addresses of the children were also recorded and comparison made with a study of the sociology department at the university by which the city of Honolulu had been mapped according to the racial origin of residents in each locality. It is thus possible to compare the effect on a child's language of his living in a section surrounded predominantly by people speaking his parental language or by those speaking a diversity of languages.

TABLE 4

Correlations Between Occupation, Language Rating and English Education of Certain Parental Groups

	Chinese *r*	Japanese *r*	Korean *r*	Hawaiian *r*	Portuguese *r*	Filipino *r*
First order correlations						
Occupation and education	.45±.048	.20±.058	.34±.059	.56±.041	.64±.041	.51±.046
Occupation and language	.55±.042	.23±.057	.30±.062	.35±.053	.43±.065	
Language and education	.74±.027	.63±.036	.71±.034	.58±.040	.74±.036	
Second order correlations						
Occupation and education (language out)	.08±.060	.07±.060	.19±.063	.47±.047	.53±.057	
Occupation and language (education out)	.36±.052	.14±.059	.09±.066	.03±.063	−.08±.080	
Language and education (occupation out)	.66±.034	.61±.038	.68±.036	.55±.041	.66±.045	

III. THE METHODS

A. MEANS OF SECURING DATA

The method of study used was similar to that used by D. McCarthy (19) and in my previous studies of monolingual children (30, 31), the results of which may, therefore, serve as norms for comparison.

For each case studied, there were recorded 50 consecutive remarks of the child, spontaneous insofar as he was never addressed by the observer. Except for two special series, the records were taken in his own home or in its vicinity and, except for the two-year-olds, when he was playing with other children. As the birth rate is still comparatively high in Hawaii except for the Caucasians, there were few homes where the child did not have brothers or sisters; in fact, there were several cases where the child had 9 to 11 older siblings. Where he had none, it was not difficult to follow him out into the yard or sidewalk where he met his playmates or to induce a neighbor's child to come over to join him. It was thought that this familiar home setting would be the best in which to study the home language of the child. The presence of other children would make a group as large as that in which most children of that age participate in free play at kindergarten or nursery school, and, therefore, records so taken would be comparable to my studies of monoglots. The more formal procedure used by McCarthy (19) was considered, but it was felt that the language used by the observer would influence *too much that used by the child, and the purpose was to learn what* was the most natural manner of speech for the children in their own homes. As it has been shown (31) that there is very little, if any, difference in the language of two-year-olds when with adults only and when with children also, it was not considered necessary to require the presence of other children when recording the speech of the youngest children. Even if the two-year-olds had brothers and sisters, if they were younger they were unable to talk, and if older they often considered the baby too small a companion; and, if the two-year-olds had no siblings, they usually had not yet acquired playmates in the neighborhood.

Twenty-five records were taken for comparative purposes when a child of three to six years was alone with adults, which series will be referred to hereafter as those taken in Situation *A*; and 100 records were taken at Castle or Moiliili kindergartens.

As the children to be studied spoke several different languages, it was obvious that no one person could secure all the records. The recorders for each racial group were, therefore, university students of the same racial ancestry as the child observed. The Chinese material was gathered in the course of work done on masters' theses by Misses Chun and Yee and on a special research project by Mrs. Motoyama. A few records were taken by students in child psychology classes, but almost all of the records other than those of Chinese children were taken by carefully selected students working under the *F. E. R. A.* or *N. Y. A.* during the school years 1934 to 1937.

Children were contacted through friends or neighbors and by calling at homes in communities where there was known to be a concentration of children of the desired race. Care was taken not to secure too many through friends, lest the sample be too select. This was rather difficult in the case of the Koreans, since their population is small. The observers met with few rebuffs, although a few cases had to be turned in anonymously and required data were occasionally refused, particularly when one parent was illegally resident in the United States.

The greatest difficulty was met with by every interviewer in securing two-year-olds. When the mothers were asked to let the recorder listen to the child talk, they sometimes refused because "baby does not talk yet," so that many of the less proficient children of that age were not observed. Our two-year-olds are, therefore, probably a somewhat select group, although some in each group were using very few comprehensible words.

The most enthusiastic coöperation was obtained by the Hawaiian observer, who was even invited to meals with the families of strangers she visited. Her records indicate a greater influence of her presence than that of any other, for many more of her cases asked her questions and appeared to be trying to be sociable with the visitor. She had no complaints to make of children too shy to talk, and next to the plantation Flipinos, the time needed to secure 50 sentences from each child averaged least of all groups. Whether this difference is due to the personality of the observer or to a greater friendliness generally believed to be found in the Hawaiian race is not known.

The time required to secure 50 sentence samples is shown in Table 5. The average for the different racial groups is fairly uni-

TABLE 5

AVERAGE TIME IN MINUTES NEEDED TO RECORD 50-SENTENCE SAMPLES ACCORDING TO AGE AND RACE

	Chinese	Japanese	Korean	Filipino		Hawaiian	Portuguese
				city	rural		
At							
2	74	61	60	55	50	50	58
3	58	41	47	44	36	37	49
4	46	46	42	48†	35	37	49
5	44	38	37	44	33	42	60*
6	35	32	37	35	29	37	52*
For							
Boys	52	40	42	47	36	45	53
Girls	49	47	48	43	37	38	53
All	51	44	45	45	37	41	53

†1 boy took two hours.

*Most of these records taken by the same person; the earlier ages by others. She was careless in time-keeping for she turned in many even hour records.

form. The Portuguese and Chinese needed the most time, due, probably, in the former case to the careless timekeeping of one observer and in the latter case to an unusual number of shy children at the two lower age levels. The rural Filipinos required the least time. Were it not that the city Filipinos were slower talkers, a race difference might be suspected. As it is, a more probable explanation is that the rural children were better acquainted with the examiner. All these records were taken by the same young man, the majority during his holidays on Maui, which was his home. He returned to the same camps each vacation period, so the children came to recognize him and hence talked more freely in his presence. He was accustomed to treat his young subjects to candy, which news was probably spread about.

There are no consistent sex differences. The averages found for each sex are invariably larger for that which included the largest number of two-year-olds, whose records took the longest time to secure; but there was only one more two-year-old girl among the Japanese, the group which shows the greatest sex difference. As there was a marked difference between the sexes at four to six years old, it is possible that the greater repression of Japanese girls has resulted in slightly reducing the fluency of their chatter.

There is a steady reduction in amount of time spent in securing the samples of speech at each age level; the one exception (Portuguese) is again probably not real but due to the same careless timekeeping.

Each observer was directed to record every word and unintelligible effort at speech used by the child until 50 such attempts had been recorded. They were to address the child at no time and to remain as inconspicuous as possible but to answer briefly if the child questioned them in order that he might not think the situation unusual. They were also to discourage the mother's attempts to urge the child to talk and to wait until the child had accepted the arrival of the visitor as a natural situation and had turned to his own affairs before beginning the record. This was usually possible by utilizing the first minutes of the interview in obtaining the necessary data from the mother.

For each child studied, the interviewer recorded the date, place, and length of time used in taking the record; attendance, if any, at school or kindergarten; and the age, sex, order of birth, address, the members of the family, the languages used by each, and the racial descent, occupation, and English education of the parents. The form used is shown in Appendix *A*.

All the recorders, except some of the Japanese, used English letters in writing and in transliterating from the other languages if necessary. The Japanese words were usually recorded by the Japanese recorders in *katakana*. For each sentence spoken by the child where the situation was not obvious from his remark, a note was made describing the circumstances under which he spoke. All non-English words were translated by the recorder if possible. In the case of sounds unintelligible to the observer but which appeared to have a definite meaning to the child, the mother was consulted to determine if such sounds were consistently used with the same meaning and were, therefore, to be considered as neologisms, regular baby words, or non-verbal expressions. When the recorder could not translate any words or syllables, they were examined by the staff of observers who spoke other languages, in the light of the interpretation placed upon them by the child's mother or as evident in the context or situation in which they were used, to determine whether they were words found in the other languages spoken in the territory, either as regular or baby words, or were merely neologisms composed by the particular child whose conversation was under examination.

More than once did recorders count as English words which had come from another language, especially the Hawaiian. In trying to trace the origin of the word *"ka"* used by some Chinese children who were jumping rope, a Caucasian stenographer born in the Islands but who spoke little but English was consulted. She averred that the word was English. She and her English-speaking playmates had always used it when jumping rope. Our Hawaiian consultant identified it as an Hawaiian word meaning "revolve" or "roll."

B. METHODS USED IN THE ANALYSIS OF DATA

It was decided to analyze all data as nearly as possible as had been done in my previous study of monoglots (30), so that comparison with that study would be possible.

One exception was made in investigating the egocentricity of speech, in which case it was decided to use the more objective method used by Fisher (11).

The first type of analysis made was according to language used. First, the number of words in each language used by each child was counted and the per cent of words from each determined. Neologisms were counted separately, and unintelligible phrases or syllables for which no meaning was ascertainable were not included in the word count.

Although there are two or three distinct dialects of Japanese, Chinese, and Filipino spoken in the Islands, it was not found feasible to classify words according to dialect of the language spoken.

Next, each sentence was examined and classified as entirely English, entirely Japanese, or Hawaiian, or any of the other languages used, or as mixed. The mixed sentences included many combinations, usually being composed of English and the parents' mother tongue; but there were all possible double combinations of the six languages spoken. There were even several sentences composed of three and a few of four languages recorded. In this classification, a sentence was not counted as mixed when it was entirely in one language, except for a proper noun or a word for which there was no equivalent in the other language or which had become sufficiently anglicized to appear in Webster's latest dictionary. For example, *"poi"* and *"lei"* were counted as Hawaiian words, but if the sentences in which they occurred were otherwise entirely in English, the sentences were not counted as mixed, since both these words are to be found in English dictionaries. If the sentence contained neologisms or baby words, they were traced to the language of origin, if ascertainable, and the sentence counted as mixed or entirely of one language accordingly. If the origin of such words could not be determined, they were considered to be corrupted forms of the language of the sentence in which they were found, and the sentence was not classified as mixed.

The count of words and classification of sentences in other languages were made by our staff of foreign-speaking students. Miss A. Keakealani Lee (16) had studied the frequency of Hawaiian words in all records, and Misses Chun (8) and Yee (36) had analyzed the Chinese words and sentences for their theses; and this material was made accessible to us. These three and all the other students who helped in the analysis conformed to the same rules. When difficulties arose, as occurred especially in classification of words by parts of speech, authorities were consulted.

Words were also classified as to the nine parts of speech, articles and adjectives being considered separately in English; but in classifying words in other languages, all types of modifiers were considered together, since a finer classification was not possible for all languages used. This classification has not been made for the Japanese words, as no one has yet been found who felt capable of the task.

Whenever it was not possible to determine the use the child made of the word, it was classified according to most common grammatical usage. The frequency of occurrence of each word, the total number of different words, and the percentage each part of speech composed of the total words of the language involved were found.

Copulas, relatives, inflected forms of words were also tabulated, and the percentages of these were found for the English portions of the records.

In determining the length of response, the rules previously used for counting words were followed. All contractions of two words into one, of whatever nature, were scored as one word. They were later arbitrarily classified under that part of speech to which the predominant part of the word belonged, but record was kept also as to the nature of the contracted element so that the frequency of use of copulas, auxiliaries, and the negative "not" could be ascertained. Hyphenated words were counted as single words, but all words written as separate words, even when combined into one verbal phrase or a double name such as "Betty Lou," were scored as separate words. The length of sentences was calculated separately for those in each language and also for mixed sentences, as well as for all types of sentences together.

Sentences were classified as was done by Fisher (11) to determine the degree of egocentricity in different categories, according to whether the subject of the sentence was the child speaking, things, or people, or whether it was non-verbal. Another category was added for those sentences which were answers to questions where the subject could not be inferred from the context. As the Japanese language infers the subject from the situation, as English does in imperative sentences, it was necessary in that language, and therefore in the other languages, to classify the sentences for this analysis according to "subject understood," wherever possible, and not as to subject stated. With these exceptions, Fisher's rules were followed.

Sentences were also classified as to form as questions, statements, imperative, and declarative, with subheads according to function as criticism, answers, naming only, social phrases, threats, and variety of imperative (which included questions such as, *"May I have it?"* and statements such as, *"I want that."*) Negative sentences were also tabulated and counted.

Another classification of the sentences was also made according to structure as simple, complex, and compound.

A special study was made of all the questions, using the adaptation of Isaacs' (15) analysis that had been used before (29).

Finally, all English sentences and mixed sentences that were at least half English were studied to determine the amount of error. Only every other mixed sentence that was exactly half English was retained for this study, so that when it should be possible to examine the sentences for error from the standpoint of the other language, such analysis could be made on an equal basis. The classification of errors was made according to usual grammatical usages. Although made in great detail at first, the headings were later grouped for convenience under such larger heads as were found to include best the errors found.

Two error indices were calculated for each child by dividing the number of errors by the number of English words used. In one case, each use of a non-English word in a sentence predominantly English was counted as an error; in the other case, these instances of mixture were not counted as errors. The introduction of a foreign proper noun or word for which there was no English equivalent or which had become sufficiently anglicized to appear in the latest edition of Webster's dictionary was considered in neither case to constitute an error.

From previous studies, certain measures of language development had been found to show less variability between individuals and more regular increase with age. When the Chinese data (the first to be completed) had been analyzed, three of these measures with the addition of per cent of English words were correlated with each other and with age to determine if they showed sufficient correlations with age to be useful in the evaluation of the various factors to be studied. The results are shown in Table 6. As the correlation of per cent of verbs conjugated with age was very low when other factors were held constant, this measure was omitted in all but the main comparisons afterward. The measures of sentence length and errors in use of English correlate +.684 and —.694 with age, when amount of English is held constant and are retained. Although they correlate quite highly with each other when amount of English

TABLE 6

Correlations of Different Measures Calculated from Chinese Data Only

	Age	Sentence length	English	Conjugated verbs
	First order correlations			
1. Age				
2. Sentence length	.714±.030			
3. Per cent English	.324±.054	.225±.057		
4. Per cent of verbs conjugated	.281±.056	.340±.053	.337±.054	
5. Errors	—.726±.029	.810±.021	—.327±.054	—.264±.056
	Second order correlations with Age held constant			
3. Per cent English		.081±.060		
4. Verbs conjugated		.207±.058	.271±.056	
5. Errors		—.605±.038	.142±.059	—.090±.060
	With sentence length held constant			
3. Per cent English	.240±.057			
4. Verbs conjugated	.060±.039		.283±.055	
5. Errors	—.357±.053		—.252±.057	.020±.060
	With amount of English constant			
2. Sentence length	.684±.032			
4. Verbs conjugated	.196±.058	.453±.048		
5. Errors	—.694±.031	—.800±.022		—.173±.059

or age is held constant, as sentence length is based on the child's use of both languages and error index on English only, it seemed desirable to continue the use of both measures. That the use of longer sentences is independent of the language spoken and, therefore, can be considered a useful measure of the progress in mastery of a verbal means of communication, irrespective of which language is used, is indicated by the correlation of only +.081 between per cent of English words and sentence length when age is held constant.

The per cent of mixed sentences, a measure also retained for all comparisons, does not have a linear correlation with age. Their absence, however, serves as a measure of the child's ability to differentiate between the two languages he is learning, except that, in the earliest stages of speech, they are lacking simply because in the stage of one-word sentences, words from two different languages could not be combined in a single sentence.

IV. RESULTS

A. The Proportion of English Used

1. *Per Cent of Words From Each Language*

The total number of words in the material gathered from all sources amounted to 181,561 words (see Table 7), of which 88.4 per cent were English, 5.9 per cent Japanese, 2.4 per cent Filipino, 1.9 per cent Chinese, and 1.1 Hawaiian. The number of Korean and Portuguese words together amounted to less than one-half of one per cent. The seven main groups averaged 86.7 per cent English words and only 12.4 per cent of the language of their ancestors other than English. Evidently English is penetrating all groups of the Islands extensively. The only other language that is used by others than those for whose parents it is the native tongue is Hawaiian; but its contribution is only 0.8 per cent. The few words from Filipino, Portuguese, and Korean used by children of other races are all proper nouns, with the exception of *"no sabe"* used once. The only Chinese word not a proper noun used by more than one non-Chinese child was the interjection *"ai-yah,"* expressive of surprise or dismay.

The use of Japanese proper nouns was sometimes accompanied by two honorifics, *"san"* and *"chan."* Besides these and the anglicized word *"kimono,"* there were only nine Japanese words used by more than one non-Japanese child. There were *"bakitale"* or crazy; *"buta"* or pig; *"daikon"* and *"saimin,"* two common Japanese foods; *"itai,"* hurts or sore; *"oppa,"* carry; *"obake,"* ghost; *"okasan,"* mother; and the most common of all, a phrase *"jun ken po,"* used in counting out in games. The syllables stand for the articles which are represented by the player's hands—"paper," "scissors," and "stone."

The Hawaiian words were counted and tabulated by Miss Lee (16). She found quite a list of words used by more than one non-Hawaiian. Only words used by at least five non-Hawaiians are listed in Table 8. They are arranged in order of decreasing numbers of different children using them.

The number of non-English words used by children whose parents did not speak the languages of origin is, then, so small that it appears that there is very little infiltration of any other language except Hawaiian, and even that is by no means extensive.

TABLE 7

TOTAL NUMBER OF WORDS IN EACH LANGUAGE USED BY THE CHILDREN IN THE DIFFERENT GROUPS STUDIED

	English	Hawaiian	Japanese	Chinese	Filipino or Spanish	Portuguese	Korean	Total
Japanese	8861	162	7873	1	1	0	0	16898
Boys	5250	80	3789	1	0			9120
Girls	3611	82	4084	0	1			7778
Chinese	16378	80	12	3542	0	0	0	19922
Boys	8511	45	5	2212				10773
Girls	7867	35	7	1240				9149
Filipino (city)	19858	303	37	5	1147†	0	1	21351
Boys	9980	166	15	3	520		1	10685
Girls	9878	137	22	2	627			10666
Rural	15062	292	50	55	3216	0	0	18675
Boys	7498	166	27	36	1664			9391
Girls	7564	126	23	19	1552			9284
Korean	18779	118	9	9	0	2	484	19401
Boys	9248	67	7	7		2	236	9567
Girls	9531	51	2	2		0	248	9834
Hawaiian	18827	606	17	7	25	1	5	19488
Boys	8743	321	8	4	11	0	1	9088
Girls	10084	285	9	3	14	1	4	10400
Portuguese	20350	107	17	1	5	155	0	20635
Boys	10118	51	12	0	3	59		10243
Girls	10232	56	5	1	2	96		10392
Rural								
Japanese	1145	47	2261	1	3			3457
Hawaiian	522	53	1	3	1			580

TABLE 7 (*continued*)

	English	Hawaiian	Japanese	Chinese	Filipino or Spanish	Portuguese	Korean	Total
Kindergarten	17662	147	229	39	4			18081
In Situation *A*	5130	55	168	3	8	21	13	5378
Haoles	17644	21*	20**	10‡				17695
Total	160218	1971	10694	3586	4410	179	503	181561
% of words from each language	88.4	1.1	5.9	1.9	2.4	0.1	0.3	
Words used by others than those whose home language it is (Seven Main Groups Only)		1062	142	73	37	3	5	
Per cent of language spoken in home that is	86.7	0.8	0.1	.06	.03	.003	.004	Home language other than English 12.4

*4 proper nouns 12 pau 5 misc.

**Mostly proper nouns.

†Spanish 83 mostly proper nouns.

‡All proper nouns except 1 Chinese exclamation.

TABLE 8

Hawaiian Words Used by Non-Hawaiian Children

pau	finished	*lei*	garland
pilau	filthy	*ka*	the
hemo	open, unfasten	*liilii*	small
make	dead	*ka*	twirl a rope
kaukau	eat	*hula*	dance
puka	hole	*poi*	food usually made from taro
hana	work	*haole*	person of Anglo-Saxon stock
hapai	lift, carry	*kanaka*	a man, usually Hawaiian
moemoe	sleep	*kapu*	forbidden
pupule	crazy	*hanapaa*	fasten
maikai	good	*poho*	lose
ono	tasty	*huhu*	angry
auwe	alas	*'olo*	lazy, awkward
kukae	filth	*makau*	fear
		mea	thing

Table 9 shows the per cent of words from each language used by each group. It will be noticed that each rural group uses words from one more language than does the city group of the same race. The kindergarten and Situation *A* groups are both mixed groups, and so naturally several languages are represented among them. Every group uses English, its own language, and Hawaiian. The *Hawaiian group, being of mixed ancestry, uses words from the* languages of all the other peoples represented in their ancestry except non-English European languages. The Filipinos use the greatest variety of languages.

2. *Per cent of English Words and Sentences and of Mixed Sentences Used by Each Racial Group*

Tables 10 and 11 show the per cent of English words and sentences used by each of the seven main groups. These per cents were calculated by averaging the per cents each child used, whereas those in Table 7 were calculated from the total number of words used by the group as a whole.

Whether it be words or sentences entirely English that are considered, the order is the same: Portuguese, Hawaiian, Korean, city Filipino, Chinese, rural Filipino, and Japanese.

Only the last three groups contain any large number of really

TABLE 9

Per Cent of Words Used by Each Group That Belong to Each Language

	English	Hawaiian	Japanese	Chinese	Filipino Spanish	Portuguese	Korean
Japanese	52.4	1.0	46.6				
Boys	57.6	0.9	41.5				
Girls	46.5	1.1	52.5				
Chinese	82.2	0.4		17.3			
Boys	79.0	0.4		20.5			
Girls	86.0	0.4		13.5			
Filipino							
City	93.0	1.4	0.2		5.4		
Boys	93.4	1.6	0.1		4.9		
Girls	92.6	1.3	0.2		5.8		
Rural	80.7	1.6	0.3	0.3	17.2		
Boys	79.8	1.8	0.3	0.4	17.7		
Girls	81.5	1.4	0.2	0.2	16.8		
Korean	96.9	0.6					2.5
Boys	96.9	0.7					2.4
Girls	96.9	0.5					2.5
Hawaiian	96.7	3.1	0.1		0.1		
Boys	96.2	3.6	0.1		0.1		
Girls	97.0	2.7	0.1		0.1		
Portugese	98.6	0.5	0.1			0.8	
Boys	98.8	0.5	0.1			0.6	
Girls	98.5	0.5	0.1			0.9	
Kindergarten	97.7	0.9	1.2	0.2			
Rural							
Japanese	33.1	1.5	65.4	0	0.1		
Hawaiian	90.0	9.1	0.1	0.5	0.1		
In Situation A	95.4	0.7	3.1			0.4	0.2
Haole	99.7	0.11	0.12	0.05			

TABLE 10

AVERAGE PER CENT OF ENGLISH WORDS AND SENTENCES USED PER CHILD AND EXTENT OF CONFUSION OF TWO LANGUAGES AS MEASURED BY THE AVERAGE PER CENT OF SENTENCES COMPOSED OF MORE THAN ONE LANGUAGE

	Average per cent								σ's of proportion at			
	2	3	4	5	6 yrs.	boys	girls	All	2	3, 4 or 5†	6 yrs.	All
Used by Japanese												
English words	48	45	45	53	61	55	45	50	10.0		9.6	4.5
English sentences	40	36	36	41	53	46	37	41	9.8		10.0	4.4
Mixed sentences	16	29	27	31	24	24	27	25	7.3	9.2	8.5	3.9
Used by Chinese												
English words	72	73	79	85	93	77	85	80	9.0		5.1	3.6
English sentences	64	69	75	82	91	73	80	76	9.6		5.7	3.8
Mixed sentences	5.1	13.2	9.8	7.6	3.8	8.7	7.3	7.9	4.4	6.7	3.8	2.4
Used by rural Filipinos												
English words	78	82	72	79	85	78	81	79	8.3		7.1	3.6
English sentences	70	64	58	62	73	63	68	65	9.2		8.9	4.3
Mixed sentences	8.9	18.5	21.9	25.2	19.4	19.4	18.1	18.8	5.7	8.9	7.9	3.7
Used by city Filipinos												
English words	84	91	94	94	96	93	91	92	7.1		3.9	2.4
English sentences	72	80	85	85	89	83	81	82	9.0		6.3	3.4
Mixed sentences	15	16	13	13	10	13.8	14.8	14.3	7.1	7.3	6.0	3.1
Used by Koreans												
English words	90	97	96	97	98	96	95	96	6.0		2.8	1.8
English sentences	78	94	91	94	95	91	90	91	8.3		4.4	2.6
Mixed sentences	3.2	2.6	6.8	4.6	3.8	4.9	3.6	4.2	3.5	5.0	3.8	1.8

TABLE 10 (*continued*)

AVERAGE PER CENT OF ENGLISH WORDS AND SENTENCES USED PER CHILD AND EXTENT OF CONFUSION OF TWO LANGUAGES AS MEASURED BY THE AVERAGE PER CENT OF SENTENCES COMPOSED OF MORE THAN ONE LANGUAGE

	Average per cent								σ's of proportion at			
	2	3	4	5	6 yrs.	boys	girls	All	2	3, 4 or 5‡	6 yrs.	All
Used by Hawaiians												
English words	93	96	97	97	96	95	97	96	5.1		3.1	1.8
English sentences	85	93	94	94	94	91	93	92	7.1		4.8	2.4
Mixed sentences	6.8	5.4	4.4	5.2	4.2	5.2	4.6	4.8	5.0		4.0	1.9
Used by Portuguese												
English words	99	99	99	98	98	98.6	98.5	99	2.0		2.8	1.0
English sentences	98	97	98	96	96	97.2	96.5	97	2.9		4.1	1.6
Mixed sentences	1.4	2.7	2.3	3.0	3.6	2.0	2.9	2.6	2.4		3.7	1.4

†This column shows σ's for the highest per cent of mixed sentences which was reached at different ages by different groups.

TABLE 10*a*

INCREASES IN THE PER CENT OF ENGLISH AND USE OF MIXED SENTENCES

	In per cent of English			In use of mixed sentences			
	Differences from 2 - 6	σ's of differences	Critical ratio	Differences		σ's of differences	Critical ratio
For Japanese							
in use of English words	13	14.0	.93	from 2 to 5 yrs.	15	12.5	1.2
in use of English sentences	13	14.0	.93	" 5 to 6 "	— 7	11.7	0.5
For Chinese							
in use of English words	21	10.3	2.0	" 2 to 3 "	8.1	8.0	1.0
in use of English sentences	27	11.2	2.4	" 3 to 6 "	— 9.4	7.7	1.2
For rural Filipino							
in use of English words	7	10.9	0.6	" 2 to 5 "	16.3	10.4	1.6
in use of English sentences	3	12.8	0.2	" 5 to 6 "	— 5.8	11.7	0.5
For city Filipino							
in use of English words	12	8.3	1.4	" 2 to 3 "	1	10.2	0.1
in use of English sentences	17	10.9	1.6	" 3 to 6 "	— 6	9.3	0.6
For Koreans							
in use of English words	8	6.6	1.2	" 2 to 4 "	3.6	6.3	0.6
in use of English sentences	17	9.4	1.8	" 4 to 6 "	— 3.0	6.1	0.5
For Hawaiians							
in use of English words	3	6.4	0.5	" 2 to 6 "	—2.6	6.4	0.4
in use of English sentences	9	8.6	1.0				
For Portuguese							
in use of English words	0.8	3.4	0.2	" 2 to 6 "	+ 2.2	4.4	0.5
in use of English sentences	2.3	5.1	0.4				

TABLE 11

DIFFERENCES BETWEEN RACIAL GROUPS AS TO PER CENT OF ENGLISH USED AND EXTENT OF CONFUSION OF LANGUAGES AS MEASURED BY THE PER CENT OF SENTENCES IN WHICH MORE THAN ONE LANGUAGE OCCURS

	Portuguese			Hawaiian			Korean		
	Diff.	σdiff.	Critical ratio	Diff.	σdiff.	Critical ratio	Diff.	σdiff.	Critical ratio
Minus Hawaiian									
English words	3	2.0	1.5						
English sentences	5	2.9	1.7						
Mixed sentences	— 2.2	2.4	0.9						
Minus Korean									
English words	3	2.0	1.5	0					
English sentences	6	3.0	2.0	1	3.5	0.3			
Mixed sentences	— 1.6	2.3	0.7	0.6	2.6	0.2			
Minus city Filipino									
English words	7	2.6	2.7	4	3.0	1.3	4	3.0	1.3
English sentences	15	3.8	3.9	10	4.2	2.4	9	4.3	2.1
Mixed sentences	—11.7	3.4	3.4	— 9.5	3.6	2.6	—10.1	3.6	2.8
Minus Chinese									
English words	19	3.7	5.1	16	4.0	4.0	16	4.0	4.0
English sentences	21	4.1	5.1	16	4.5	3.6	15	4.6	3.3
Mixed sentences	— 5.3	2.8	1.9	— 3.1	3.1	1.0	— 3.7	3.0	1.2
Minus rural Filipino									
English words	20	3.8	5.3	17	4.0	4.2	17	4.0	4.2
English sentences	32	4.5	7.1	27	4.9	5.5	26	5.0	5.2
Mixed sentences	—16.2	4.0	4.0	—14.2	4.2	3.4	—14.6	4.1	3.6
Minus Japanese									
English words	49	4.6	10.6	46	4.8	9.6	46	4.8	9.6
English sentences	56	4.7	11.9	51	5.0	10.2	50	5.1	9.8
Mixed sentences	—22	4.1	5.4	—20	4.3	4.6	—21	4.3	4.9

TABLE 11 (*continued*)

DIFFERENCES BETWEEN RACIAL GROUPS AS TO PER CENT OF ENGLISH USED AND EXTENT OF CONFUSION OF LANGUAGES AS MEASURED BY THE PER CENT OF SENTENCES IN WHICH MORE THAN ONE LANGUAGE OCCURS

	City Filipino			Chinese			Rural Filipino		
	Diff.	σdiff.	Critical ratio	Diff.	σdiff.	Critical ratio	Diff.	σdiff.	Critical ratio
Minus Chinese									
English words	12	4.3	2.8						
English sentences	6	5.1	1.2						
Mixed sentences	6.4	4.0	1.6						
Minus rural Filipino									
English words	15	4.4	3.0	1	5.1	0.2			
English sentences	17	5.5	3.1	11	5.7	1.9			
Mixed sentences	— 4.5	4.8	0.9	—10.9	4.4	2.5			
Minus Japanese									
English words	42	5.1	8.2	30	4.6	3.7	29	5.8	5.0
English sentences	41	5.6	7.3	35	5.8	6.0	24	6.1	3.9
Mixed sentences	—11	5.0	2.2	—17	4.6	3.7	— 6	5.4	1.1

bilingual children, for the per cent of English words used is above 90 for the other four. All groups, except the Portuguese, show an increase in the use of English from two to six years. The curve of increase is shown in Figure 4.

The Portuguese curve is almost perfectly horizontal. Their slight decrease of 0.8 in English usage is due to a very slight increase in

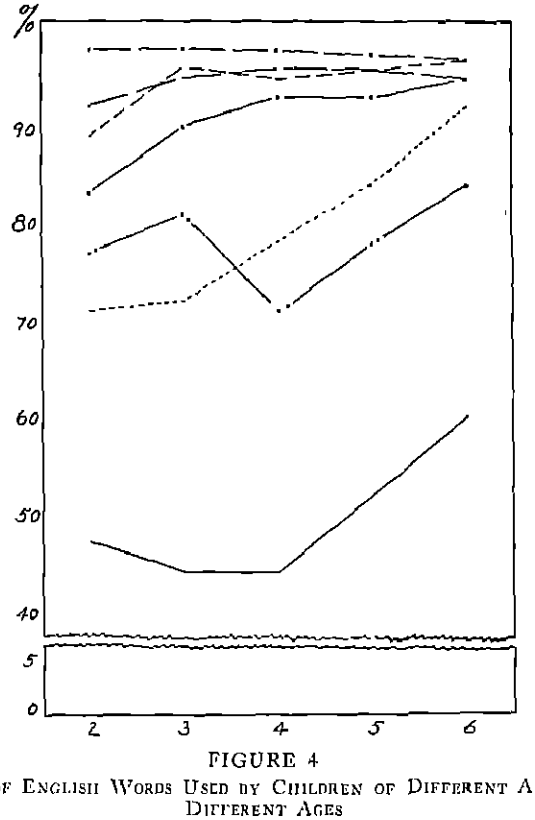


FIGURE 4

PER CENT OF ENGLISH WORDS USED BY CHILDREN OF DIFFERENT ANCESTRY AT DIFFERENT AGES

Portuguese
Hawaiian
Korean
City Filipino
Rural Filipino
Chinese
Japanese

Hawaiian and, on the part of the girls, to an increase in Portuguese. Probably their parents use English almost entirely when speaking to their children who, as they grow older and are more often away from home, pick up a few Hawaiian words from their playmates and a few Portuguese words from their grandparents. They do not learn enough of any other language to frame complete sentences, so that their use of non-English words occurs almost entirely in combination with English words in mixed sentences.

Although our raters usually marked English as the preferred language in Hawaiian homes and Korean preferred in Korean homes, the per cent of English words used by the groups is the same and that of sentences differs by but one per cent for these groups; but the Korean children use less English at two years and more at six than do the Hawaiian. Evidently, their parents do use Korean more, but the children, from their older brothers and sisters and through their contacts with playmates of different linguistic background, quickly add English to their vocabularies and spurn the use of Korean. Even in their preschool years, these children and their parents must find difficulty in understanding each other, for many conversations, where it was recorded that the parents spoke very little English, are, nevertheless, mainly in English.

The per cent of mixed sentences used by these children shows the rise between two and six years typical of bilingual groups, but by six years, the per cent is not much more than half that found at four years and is almost the same as for the six-year-old Portuguese (see Figure 5). At that age, only one per cent of Korean sentences occur. The Korean words used are mainly the words for members of the family and household articles which they do not often use except at home and for which their opportunity of learning English names is slight. So they must continue to use these few words for which they know no English. But not all of their mixed sentences are thus composed. They have also picked up a few Hawaiian words; and, as all the groups studied. do at times insert them in English sentences. Quite probably many of the children who do so are under the same impression as the *haole* woman mentioned above, who thought "*ka*" was an English word, and do not know that they are mixing two languages in one sentence.

The Hawaiian children show a steady decline in mixed sentences,

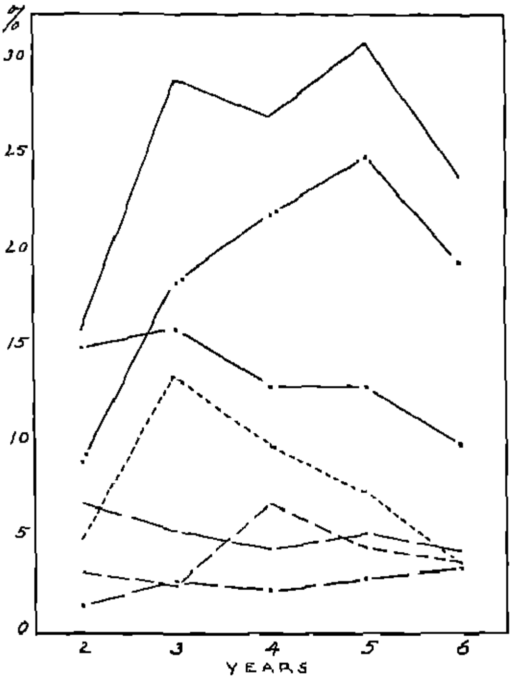


FIGURE 5

PER CENT OF SENTENCES COMPOSED OF MORE THAN ONE LANGUAGE USED BY CHILDREN OF DIFFERENT ANCESTRY AT DIFFERENT AGES

Japanese ————————
City Filipino . ——————— .
Rural Filipino . . —————— . .
Hawaiian ——— ———
Chinese - - - - - - - -
Korean — — — — —
Portuguese . ——— ——— .

but they confuse the two languages still at six years. Their mixed sentences show their mixed heritage, for they include words from several other languages. Twice as many of the six-year-old's sentences are entirely in their ancestral language than is true of the Koreans; but the amount is only 2.1 per cent. As English is the preferred language of their parents, there is no difficulty in understanding each other; but the meager amount of Hawaiian used shows that that language is fast dying out. No difference between the three above groups is statistically significant.

The Filipino children in Honolulu use over 90 per cent English words, but they and their rural brothers and sisters show the greatest confusion of language. Their mixed sentences are the result of many different combinations of languages up to as many as four varieties occurring in one sentence. One of these combinations was Filipino, English, Chinese, and Hawaiian; another, the same with Japanese substituted for Hawaiian. Twenty-two different combinations were found among the rural samples; 11 among the city samples. Yet the few Spanish words, mainly proper nouns (because most of the Filipino children bore Spanish names) and the three dialects represented were all counted under the head, Filipino. It was fortunate that our Filipino observers were as good linguists as the children.

The rural children who use four times as much Filipino as the city children reach the peak of mixed sentences two years later. Except for mixed sentences, the differences between these two groups are statistically significant. The rural group uses a significantly lower percentage of English words and sentences than do four of the other groups. The two very small samples of rural Hawaiian and rural Japanese both used less English than did their city fellows.

The Chinese average slightly more English words and a few more English sentences than do the rural Filipinos. They use less English at two years but more at six years. More of their homes were reported to use mainly Chinese than were Filipino homes preferring the use of a Filipino dialect. The greater number of contacts in the city with children speaking more English, and the fact that so many of the Chinese group attend kindergarten at an early age, result in much greater progress in learning English. This group makes more progress in the use of English than does any other. This is the only group where the difference from two to six years is twice or more its standard error. The per cent of mixed sentences for this group reaches its peak at three years. Most of the sentences are a combination of Chinese and English. Some Hawaiian is mixed in at every age but no other language. They have probably differentiated pretty well between Chinese and English, not so well between English and Hawaiian, by six years, since the per cent of mixed sentences is so low as probably to include almost entirely sentences where the foreign word is inserted because the English word is unknown and not because of inability to differentiate between the two.

The Japanese use the most of their parents' language of any group, and they mix the two languages most. The differences between them and the other groups are all statistically significant except two—the proportion of mixed sentences in comparison with the Filipinos. English words and sentences show no appreciable increase until the rise from five to six years. Mixed sentences rise in number at three years and then fall off but very little, so that at six years almost a quarter of their sentences are combinations of more than one language. The number of English words is slightly less and of sentences entirely English only a little more than the Japanese until five years when, with the start of attendance at kindergarten and school, English begins at last to show a real increase. The fact that, while English words are less numerous than Japanese, the opposite is true of sentences, is due to the presence of many sentences that are Japanese except for the one pronoun *"me."*

3. *Conclusions from Data as to Proportion of English and Mixed Sentences*

To conclude the analysis of proportion of English, we find that English is distinctly the predominant language of the preschool group in Hawaii: for, except in rural districts and in the case of the Japanese children, by the time children are six years old, they use over 90 per cent of English words when playing even about their own homes. With the exception of the Japanese and Filipinos, they have learned to distinguish reasonably well between the languages current in the Islands, and over 95 per cent of their sentences are composed of but one language.

The Portuguese are really monolingual, and the Hawaiians and Koreans nearly so, when babyhood is past. Every group of children, unless it be the Japanese, prefers to use English when at play if able to, and does so in the majority of cases. This is true even in the case of the Koreans, whose parents prefer to use Korean.

Hawaiian is the only language besides English that has interpentrated the others to an extent great enough to influence the children's speech, but it composes only a small per cent of even the Hawaiian children's speech.

Mixed sentences show a different curve for almost every group. They do not occur while the child is in the stage of one-word sentences and are, therefore, comparatively rare in the earlier stages of speech. They increase in number as soon as the child combines words to any extent if he hears more than one language spoken, until he not only is able to differentiate between the languages he hears but also has acquired enough of a vocabulary in both so that he can express himself entirely in the language he wishes to use. The decrease in the number of mixed sentences from three to four years found in two groups and the smaller decrease from two to three years found in the Hawaiian data suggest that children as young as three years are able under favorable circumstances to differentiate between the languages they hear. The failure of any group to eliminate such sentences entirely indicates the difficulty of acquiring adequate vocabularies in two languages at an early age. The occurrence of mixed sentences composed of English and another language not that of the child's parents suggests a confusion of such languages with English when the source of both come from playmates.

The delay in reaching a peak in mixed sentences in the case of the rural and Japanese children may be due to the difficulty in acquiring adequate English vocabularies which are needed in speaking to playmates when parents are unable to assist the child. The decrease in such sentences found in six of the seven groups from five to six years is doubtless due to the fact of school entrance, with the twofold result of acquiring a more adequate English vocabulary and teacher's aid in differentiating between languages.

4. *Sex Differences as to Proportion of English*

Sex differences in the proportion of English used are not marked and are not significant in any group. The largest differences are in the case of the Japanese, where boys use most, and the Chinese, where girls do. In the case of the Japanese, the girls at two years use 49 per cent of English words to the boys' 47 per cent. After that, the boys use most, the difference being especially marked at four and six years. Mr. Miyamoto (22), in his study of Japanese as used in the local language schools, found the girls to excel in the use of Japanese. As girls are kept more closely at home than are their brothers, their opportunity to pick up English from playmates is less and to acquire better Japanese from their parents is greater.

The Chinese girls use more English than do boys at every age except at four years, when the boys use 80 per cent English and the girls, 78 per cent. No explanation for this difference is found.

B. ERRORS IN THE USE OF ENGLISH

1. *Frequency of Error*

To measure the amount of error found in the use of English, the number of errors each child made was counted and divided by the number of English words used. Where a sentence was only partly in English, it was included in the analysis of error if the English words exceeded those of any other language used in that sentence. If the numbers of words in two languages were equal, than alternated sentences were taken. If the sentences were so confused that more than one method of correction was possible, that one was used which would credit the child with the smallest number of errors.

Each insertion of a word from another language was counted as an error, provided such words were not proper nouns, had English equivalents, and were not anglicized to the extent of appearing in an English dictionary. Error indices were calculated with and without the inclusion of these errors of mixture. Careful study of these two measures showed that the truer picture of facility of English speech was shown by that index which included the errors of mixture, so this measure is the one used in most comparisons.

Table 12 shows the average error index and standard deviation at each age level and for the different sexes, and all children for each of the seven major groups studied.

Only for the Japanese does a sex difference of any size appear. For them, the girls exceed the boys in amount of error by an amount that is 2.3 times its standard error, or enough so that the chances are 99 in 100 that the difference is a true one. The other differences are all very small, but in every case favor the girls.

Every group shows a significant decrease of error from two to three years, and after that a further decrease which is not great enough to amount to statistical significance, even if all the next three years are considered, except in the case of the Chinese and

TABLE 12

Comparison of Error Indices of Different Races

	Age								Differences of averages					
	2	3	4	5	6	All	Boys	Girls	boy-girl	2-3	3-4	4-6	3-6	2-5 yrs.
Chinese														
av. error index	709	389	332	320	290	407	409	404						
σdist.	146	88	100	92	93	188	196	176						
σav.	29	18	20	18	18	16.7	24	24						
Diff.									5	320	59	42	99	389
σdiff.									34	32.6	26.9	27.0	25.5	32.6
D ÷ σ									0.14	9.8	2.1	1.6	3.9	11.9
Portuguese														
av. error index	501	372	376	314	329	379	386	371						
σdist.	154	156	121	95	62	139	129	150						
σav.	31	31	24	19	12	12.5	16	19						
Diff.									15	129	— 4	47	43	187
σdiff.									25	43.6	39.2	26.8	33.2	36.4
D ÷ σ									0.6	3.0	0.1	1.8	1.3	5.0
Hawaiian														
av. error index	653	456	327	314	373	425	439	412						
σdist.	185	155	137	99	73	185	180	185						
σav.	37	31	27	19	15	14.8	23	23						
Diff.									27	197	129	—46	83	339
σdiff.									32	48	41	31	34	42
D ÷ σ									0.8	4.1	3.1	1.5	2.4	8.1
Korean														
av. error index	628	464	425	385	364	453	465	440						
σdist.	170	163	77	106	57	155	144	165						
σav.	34	33	15	21	11	13.8	18	21						
Diff.									25	164	41	59	100	243
σdiff.									28	47	35	19	35	40
D ÷ σ									0.9	3.4	1.2	3.1	2.9	6.1

TABLE 12 (*continued*)
COMPARISON OF ERROR INDICES OF DIFFERENT RACES

	Age								Differences of averages					
	2	3	4	5	6	All	Boys	Girls	boy-girl	2-3	3-4	4-6	3-6	2-5 yrs
Filipino														
City														
av. error index	667	480	437	402	403	478	482	473						
σdist.	118	104	74	95	81	138	144	131						
σav.	24	21	15	19	16	12.3	19	16						
Diff.									9	187	43	34	77	265
σdiff.									25	32	26	22	26	31
D $\div$ σ									0.4	5.9	1.4	1.5	3.0	8.6
Rural														
av. error index	780	604	538	540	472	587	594	579						
σdist	141	155	118	119	53	161	156	167						
σav.	28	31	24	24	11	14.5	19	22						
Diff.									15	176	66	66	132	240
σdiff.									29	42	39	26	33	37
D $\div$ σ									0.5	4.2	1.7	2.5	4.0	6.5
Japanese														
av. error index	745	558	465	486	448	540	516	598						
σdist.	157	125	183	164	142	147	184	198						
σav	31	25	37	33	28	19.2	23	26						
Diff.									—82	187	93	—17	110	259
σdiff.									35	40	45	46	38	45
D $\div$ σ									2.3	4.7	2.1	0.4	2.9	5.8

Filipinos. The Hawaiian decrease from three to five years is significant, and the apparent failure to continue to improve may be due to a difference in sampling at six years, for the fathers of children at that age had a lower average Barr rating. It is to be noted, however, that the other non-bilingual group, the Portuguese, also shows a higher average error index at six years. It is possible that with increased age, involving a greater contact with other language groups who speak less English and speak it less well, the English of these groups is becoming contaminated.

That this may be the case is suggested also by a comparison of the differences from five to six years with the average score for all ages. The average difference from five to six for the three groups with the lowest average index for all ages is + 15, or an increase. The two middle groups, Korean and city Filipinos, show an average decrease of 10 points, and the two poorest groups show a decided improvement of 53 points. So it seems that as the children become older and mingle more together outside the home, the differences between the groups are smoothed out and their status becomes more uniform.

The average error index at two years old is, in all but one case, higher than that of 580 found for 65 monolingual two-year-olds (28). The exception is in the case of the Portuguese, who are not bilingual. If the error indices calculated without including errors of mixture (see Figures 6 and 7) are compared instead, they, too, are found to exceed the group previously studied, except again in the case of the Portuguese. This suggests a greater difficulty for bilingual children in the very beginning of their speech development.

The drop in number of errors from two to three years is in line with that found previously, but for no group was the decrease at all comparable, and even at six years, the lowest average, 290, was eighty points above that found in the monoglot series at three years. How much of the slower improvement is due to pidgin English and how much to the difficulty of mastering two languages is hard to determine.

Our groups are probably exposed to a similar degree to pidgin English. Comparing, therefore, the four groups who were found to use but little of another language, since over 90 per cent of their words were English, with the three truly bilingual groups, we find

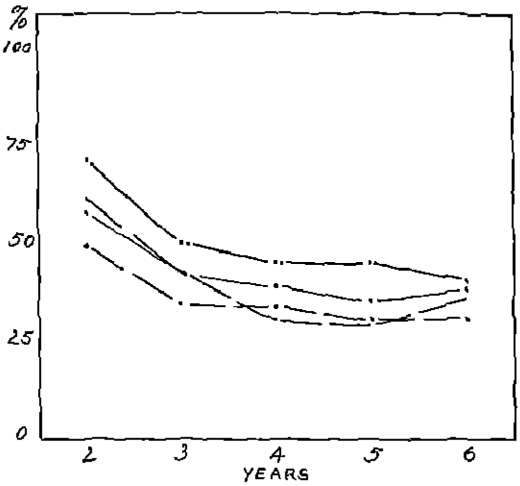


FIGURE 6

AVERAGE ERROR INDEX EXCLUDING ERRORS OF MIXING FOR CHILDREN OF DIFFERENT ANCESTRY AT DIFFERENT AGES

City Filipino . ———— .
Rural Filipino . . ———— . .
Hawaiian ——— ———
Portuguese . ——— ——— .

the combined average of error of the former to amount to 434 per one thousand words, while that of the latter is 507 per thousand, a difference of 73.

The different racial groups follow much the same order (Table 13) as they did in per cent of English used, the exception being that the rural Filipino and Japanese change places and the Chinese rise from fifth to second place. The two groups who rise in rank have the largest number in attendance at kindergarten and nursery school, the Chinese ranking first and the Japanese second.

Whereas the kindergarten attendants are almost exclusively at the six-year level for the other groups, in the case of the Japanese, many at the five-year level had just begun such attendance, while the Chinese attendants included even two- and three-year-olds.

2. *Types of Errors*

The main subdivisions of the larger groups of errors are shown

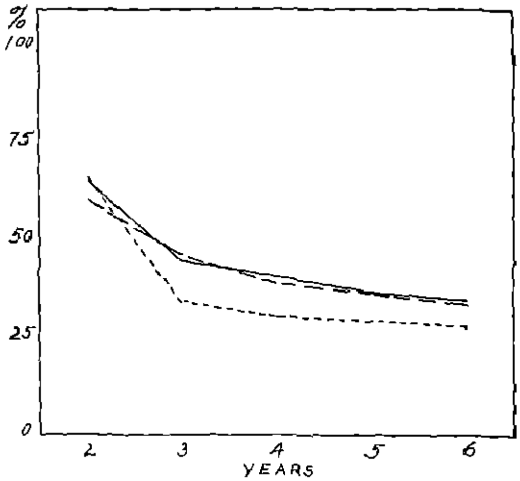


FIGURE 7
Average Error Index Excluding Errors of Mixing for Children of Different Ancestry at Different Ages

Japanese ————————
Chinese - - - - - - - - -
Korean — — — — —

in Tables 14 and 15; other tables showing additional detail are given in Appendix C.

The highest number of errors is due to the omission of essential words that leave the sentences incomplete. The omission of verbs, subjects or expletives, objects, and copulas were the the only cases where omission was considered as leaving the sentences essentially incomplete. Such error is frequent in immature speech and is the only type which was found to decrease regularly with age for all groups. It is a distinctly immature type of error and indicates a real failure on the part of the child to gain control of speech. It, probably more than any other type, measures the bilingual handicap in contrast to the pidgin English. It will be noted that this error type is more frequent in the three bilingual groups than in the four others. However, there is one type of incomplete sentence, that where the copula is omitted, which is less frequent for the bilingual groups. In pidgin English, this element is often omitted or a verb such as "*stay*" is substituted for it.

TABLE 13

DIFFERENCES BETWEEN RACIAL GROUPS IN ERROR INDICES

			Difference in size of error index	σ dist.	σ av.
Rural Filipino	—	Japanese	47	24.9	1.9
" "	—	City Filipino	109	20.0	5.4
" "	—	Korean	134	21.0	6.4
" "	—	Hawaiian	162	21.7	7.5
" "	—	Chinese	180	23.0	7.8
" "	—	Portuguese	208	20.2	10.0
Japanese	—	City Filipino	62	22.8	2.7
"	—	Korean	87	21.4	4.1
"	—	Hawaiian	115	24.2	4.8
"	—	Chinese	133	25.4	5.2
"	—	Portuguese	161	22.9	7.0
City Filipino	—	Korean	25	18.5	1.4
" "	—	Hawaiian	53	19.2	2.4
" "	—	Chinese	71	20.7	3.4
" "	—	Portuguese	99	17.5	5.6
Korean	—	Hawaiian	28	20.2	1.4
"	—	Chinese	46	21.7	2.1
"	—	Portuguese	74	18.6	4.0
Hawaiian	—	Chinese	18	22.3	0.8
"	—	Portuguese	46	19.4	2.4
Chinese	—	Portuguese	28	20.9	1.3

The omission of subject is allowable more often in Korean and Japanese than in English, but these groups do not show a higher frequency of this error type than do the others, the high frequency of errors of omission by the Japanese being due to an excessive number of very fragmentary sentences lacking verbs or verb and subject both.

For the Japanese groups, the next most frequent error is that of inserting foreign words, while for all the other groups, errors in the use of the verb rank second. Errors of mixture show very little decrease with age until five years for the group as a whole (see Figure 8). It will be recalled that mixed sentences show a different age curve for each group; but after five years, there is a drop of 33 per cent in such errors coincident with the time the children start to kindergarten and school. Although over half of

TABLE 14

TYPES OF ERRORS MADE BY ALL SEVEN GROUPS AT DIFFERENT AGES

	2 yrs.	3 yrs.	4 yrs.	5 yrs.	6 yrs.
Per cent of:					
Incomplete sentences	57.6	36.9	32.9	28.3	30.7
Verbs	12.9	21.3	22.4	26.3	24.9
Substantives	4.9	7.1	7.0	7.0	6.7
Prepositions and infinitives	4.7	8.4	10.0	10.9	10.6
Agreement	0.9	1.3	1.0	1.7	1.9
Negatives	3.6	6.3	6.6	6.4	6.3
Articles	5.9	5.4	5.7	4.4	5.0
Miscellaneous	4.4	5.3	6.6	6.0	7.6
Mixed-in foreign language	6.4	8.1	7.9	8.9	6.0
Average error index	669	475	442	394	382
Boys	665	490	446	407	378
Girls	672	462	416	383	384†
Average error index exclusive of mixed language	625	428	379	353	352
Boys	621	439	388	358	355
Girls	634	415	376	346	347
Error per 1000 words					
Total	623.4	446.0	397.1	380.4	371.1
Incomplete	365.3	167.0	124.7	106.6	111.9
Verbs	69.9	93.7	94.1	98.3	92.0
Substantives	28.4	32.7	28.9	27.3	25.7
Prepositions and infinitives	28.1	36.0	38.9	39.1	39.3
Agreement	5.7	5.9	3.4	6.1	6.4
Negatives	22.1	25.9	24.4	22.7	23.3
Articles	35.4	23.4	21.7	17.0	18.6
Miscellaneous	28.0	22.7	27.4	25.9	29.3
Mixed-in foreign language	41.0	39.6	33.7	37.0	24.7

†Japanese girls exceptionally high.

the kindergarten attendants were Japanese, the errors of mixture made by them are as few as those made by the Portuguese group. Evidently kindergarten attendance helps the child in learning the English words necessary to compose sentences entirely of that language and in identifying the origin of words.

The Filipino children make many errors of mixture. They insert in their English sentences not only words from the different Filipino dialects spoken in the Territory, but also a very few Spanish words, and they use more Hawaiian and Japanese words than does any group whose parents are not of such ancestry. They seem to be more confused than any other group in recognizing

TABLE 15

ERRORS PER 1000 ENGLISH WORDS MADE BY DIFFERENT GROUPS OF CHILDREN

				Filipino				Japanese			
	Caucasian	Portuguese	Hawaiian	Honolulu	Plantation	Korean	Chinese	City	Rural	Kindergarten	In situation A
Incomplete sentences	55	134	137	128	171	145	136	197	163	123	71
Verbs	17	85	99	106	121	102	79	44	38	73	40
Substantives	2	14	21	35	32	24	15	56	88	19	7
Articles	7	26	17	22	21	28	17	22	26	37	9
Prepositions (including sign of the infinitive)	4	36	45	40	42	43	33	24	20	28	24
Negative	0.4	22	16	27	35	24	22	21	15	18	6
Agreement	4	7	5	7	5	6	7	3	0	4	7
Order	—	7	4	9	11	14	2	5	24	4	6
Miscellaneous	5	20	12	30	35	13	15	13	20	31	11
Other language mixed in	0	8	15	39	64	12	19	80	167	8	12
Total	94	360	371	443	536	411	345	464	562	336	193

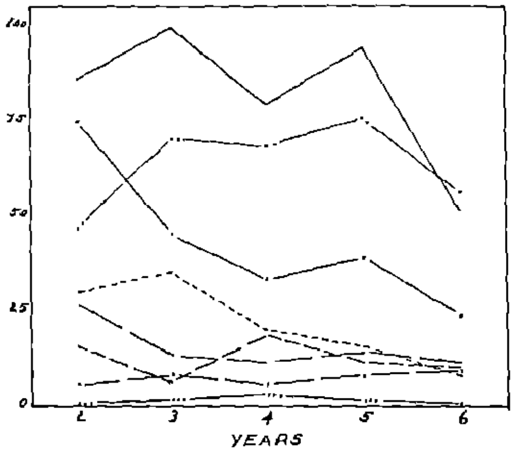


FIGURE 8

Average Number of Errors Made by Children of Different Ages and Ancestry Through Inserting a Non-English Word in an English Sentence

Japanese ————————
City Filipino . ——————— .
Rural Filipino . . —————— . .
Chinese – – – – – – – – –
Hawaiian ———— ————
Korean — — — — — —
Portuguese . ———— ———— .
Kindergarten . . . ————— . . .

English words. Probably their parents are too, for they are the most recent arrivals in the Islands, and knowing English is the language supposedly spoken in Hawaii, they are inclined to think any non-Filipino word used in speaking to them is English. Many times a word new to the observers was assumed by them to be English. This assumption must often be made by the less well-educated population.

Errors of mixture are a type of error that can be considered due to bilingualism more than to pidgin English. The failure to eliminate such errors before school entrance indicates that the young child needs to be in an unusually favorable situation if he is to differentiate between two languages. It is a peculiarly unfortunate error, as it becomes impossible for anyone not knowing both languages to understand the speaker. (See first three records in Appendix *B.*)

The most frequent difficulty in the use of verbs is in the omission of some essential element of the verbal phrase. Errors in the formation of a past tense are seldom the ones found frequently among the Caucasian group, i.e., the using of regular verb endings on irregular verbs. They are almost always the use of the present form, either alone or by preceding it by *"go"* or, less often, by *"been," "gone"* or *"went,"* as in *"I been buy the doll," "he go hit me,"* and *"I went go,"* or by addition of the adverb *"already."*

The future tense is about half the time formed by the use of *"go,"* sometimes in conjunction with a much-slurred *"by-and-by,"* as in *"by'n by he go make."* Other times it was only from the context that it could be determined that the verb in the present tense had been used with a future reference. Occasionally it is formed by the addition of *"yet."*

The most frequent wrong choices of verbs are due to an overuse of the verbs *"make,"* in the sense of *"do"* or *"fix;" "like"* where *"want"* would have been preferred by mainland youngsters; and *"lend"* for *"let."*

Errors in verb forms are less numerous at two years, owing to their very frequent omission; and for the Japanese children, because a Japanese verb is often used.

Most of these verb errors are those noted by Reinecke (26) as characteristic of the local dialect; hence, it is to be expected that the children should make such errors. That verbs are a source of difficulty for children not exposed to corrupt English is evident from the fact that they rank second in amount of error when used by the Caucasian group. Moreover, as pointed out above, none of the home languages have conjugations exactly corresponding to the English.

The high frequency of error in the use of substantives by the Japanese is due to their use of English pronouns in preference to the longer Japanese forms; also, they seem as a rule to have learned only the pronoun *"me"* and so use it always for subject or object and attach the Japanese possessive sign to it when *"my"* should be used. Errors in number are usually of singular for plural noun or of the pronoun *"them"* for *"it."* As this word is usually pronounced *"em"* or *"um,"* it may be a corruption of *"him,"* but all doubtful cases were classed as error of number. Substantive errors decrease very slightly from five to six years.

Most of the errors of gender are in using masculine pronouns for feminine; the others are almost all due to the use of "*he*" for "*it*."

The Portuguese, whose use of gender for pronouns and number for substantives most nearly corresponds to English, and the Japanese, who use so many elliptical sentences and so much Japanese that the chance of other errors is diminished, make the fewest errors of these two types.

Correct use of articles shows improvement with age from two to five years. These errors are of three kinds: omission, and substitution of either "*one*," or, occasionally, "*the*," for the indefinite article. These errors are of high frequency, but referring above again, we find that Chinese, Japanese, and Korean lack articles altogether, and although Portuguese, Filipino, and Hawaiian make use of both definite and indefinite articles, the indefinite article is the same as for the numeral *one* in Portuguese and Filipino. The Hawaiians, who have a separate indefinite article, tie with the Chinese for the lowest number of errors in the use of the article.

Errors in the use of the preposition and infinitive are, with rare exceptions, due to omission. These errors increase slightly at first, due to the more frequent need for prepositions as the child expresses more complex ideas. Thereafter, their frequency remains constant.

The Oriental languages and Hawaiian have no infinitive. The children from these linguistic groups all make more errors by omission of the infinitive sign than in the use of regular prepositions, while the reverse is true of the Filipinos, whose language possesses an infinitive. Portuguese also has an infinitive, and the difference between the frequency of these two error types is negligible for the children of that ancestry.

Errors in agreement occur less frequently when a large number of mixed sentences is used, for the rules of agreement are not the same in different languages; so if the noun was in Japanese and the verb in English, no attempt was made to discover if they might be in disagreement, and only an error of mixture was checked. These errors increase with age, partly due to the dropping out of mixed sentences, partly to the increasingly complex ideas to be expressed.

Errors in the use of the negative remain constant at all ages. Pidgin English uses "*no*" or "*never*" as the only negatives, and so nost of our subjects. A double negative is used very seldom. ; use of *"no"* for *"not"* is found among the very youngest Cauıns and would be a natural consequence for children speaking of the languages involved in the present study because of the that there is no such differentiation in the foreign language, : in all five, the same word is used in answering a question negay, as in modifying a verb and, in some cases, even in modifying ıbstantive. The Japanese, who have a word *"iie"* equivalent etimes to the English *"yes,"* sometimes to the English *"no,"* e very slightly fewer errors in using *"no"* for *"not."*

rrors in word order are so few for the Caucasians that they were d under "miscellaneous." They are sometimes difficult to detect nixed sentences. The Chinese, Hawaiian, and Portuguese make fewest errors of order. Chinese word order is quite similar English, the biggest difference being that adverbs precede verbs . pp. 617, 618). Portuguese also has a similar order. The r languages do not. As the Hawaiian children speak little but lish, the different Hawaiian order probably does not affect their ch much. The Japanese show the highest number of these rs.

nder "miscellaneous" are grouped the remaining errors: those he use of modifiers and connectives, usually of omission; errors omparison of adverbs and adjectives; redundancies; confusion of s of speech with each other; and a few not classifiable under any he above headings.

rrors in the use of connectives are most numerous among the ıps using longer sentences and are merely an accompaniment of ıttempt to use more complex sentences. To a certain extent this ue of the use of modifiers. Using one part of speech for another nore frequent for the Japanese and Koreans, redundancies for Filipinos. The fact that Filipino dialects require ligature and rate words to express plurals which would be quite superfluous English may be partly responsible for the high use of redundant ds among the Filipinos.

3. *Special Idiomatic Errors*

'ertain specific errors of high frequency or unusual picturesquewere selected for special study along with certain common

TABLE 16

LIST OF SPECIFIC ERRORS SELECTED FOR SPECIAL STUDY

	Number of children making each error								
	Filipino								
	Honolulu	plantation	Japanese	Chinese	Korean	Portuguese	Hawaiian	Kindergarten	Total
already*	10	12	7	11	2	7	5	3	57
all same—just like	4	1	0	1	0	10	0	0	16
as why—because	11	9	0	4	14	11	5	3	57
been buy—bought	6	3	2	4	1	7	5	0	28
break—tear	4	2	0	3	7	1	3	0	20
broke—break	7	16	12	10	8	7	3	1	64
"bring it come"	1	3	0	2	0	0	0	0	6
Chance—turn	0	0	2	4	5	14	2	2	29
everytime—always	4	1	4	3	4	9	9	1	35
find—hunt	1	0	0	2	0	0	0	0	3
for as infinitive sign	0	0	0	2	0	3	2	3	10
funny kind—queer	7	5	7	5	4	3	0	5	36
get—have, is, etc.	69	58	12	45	53	49	64	18	368
go as an auxiliary	88	87	30	40	81	51	26	25	428
house—family	0	0	3	0	0	0	0	2‡	5‡
kind used redundantly	44	25	35	31	27	25	11	8	206
lazy—am tired of	0	0	1	1	3	2	4	4	15
lend—let	14	3	11	12	5	1	8	2	56
like—want	77	56	44	61	77	47	91	5	458
little more—soon almost	2	8	4	1	5	4	3	3	30
long time—for a long time ago	1	1	0	4	0	2	0	3	11
make—do, fix	42	54	15	19	45	35	13	15	238
me, I—I†	3	2	3	0	4	1	2	14	29
more big	12	5	4	3	5	26	6	6	67
more bigger	4	4	3	3	3	26	5	2	50

TABLE 16 (*continued*)

LIST OF SPECIFIC ERRORS SELECTED FOR SPECIAL STUDY

	Number of children making each error							
no—not	119	113	48	98	112	108	103	21	722
no more	46	54	21	11	48	36	28	8	252
no need	2	3	3	3	5	6	3	3	28
one—a, an	16	6	4	16	17	33	12	4	108
open—turn on, unfasten	0	0	1	2	4	0	8	0	15
plenty—many	10	8	6	12	15	20	15	7	93
scare—be afraid	8	3	6	8	3	5	1	0	34
shame—bashful, be ashamed	4	4	2	4	2	0	4	0	20
sore—hurt	3	2	2	3	0	2	11	4	27
stay—is	37	32	3	9	21	37	9	10	158
"take it go"	4	2	0	3	2	1	2	0	14
tell—say, ask	15	22	0	6	21	15	2	2	83
the—how	6	5	4	3	5	7	1	6	37
too—very	1	1	2	0	2	7	4	0	17
try as auxiliary	15	16	11	15	18	16	2	4	97
waste time—don't want to	17	6	14	7	18	6	13	7	88
went go	26	24	3	7	49	18	9	8	144
wild—scold, angry	2	6	1	0	0	9	5	0	23

[*] Used sometimes to express simple past time, sometimes redundantly merely for emphasis.

[†] Only ones using this were Japanese.

[‡] Only one non-Japanese used it before five years old.

TABLE 16*a*

List of Specific Errors Selected for Special Study

	Weighted frequency of each error						
	Filipino	Japanese	Chinese	Korean	Portuguese	Hawaiian	Kinder-garten
multiplied to give all races equal weight	†	x42	x24	x20	x19	x20	x40
already	245	294	264	40	133	100	120
all same—just like	50^{H}	0	24	0	190	0	0
as why—because	217	0	96	280	209	100	120
been buy—bought	95	84	96	20	133	100	0
break—tear	63	0	72	140	19	60	0
broke—break	266^{P}	504	240	160	133	60	20
"bring it come"	47^{P}	0	48	0	0	0	0
chance—turn	0	84	96	100	266	40	80
everytime—always	50^{H}	168	72	80	171	180	48
find—hunt	10	0	48	0	0	0	0
for as infinitive sign	0	0	48	0	57	0	120
funny kind—queer	129	294	120	80	57	0	200
get—have, is, etc.	1380	504	1080	1060	931	1280	720
go as an auxiliary	1924	1260	960	1620	969	520	1000
house—family	0	126	0	0	0	0	80
kind used redundantly	730	1470	744	540	475	220	320
lazy—am tired of	0	42	24	60	38	80	160
lend—let	170^{H}	462	288	100	19	160	80
like—want	1433	1848	1464	1540	893	3820	200
little more—soon almost	119^{P}	168	24	100	76	60	120
long time—long ago for a long time	22	0	96	0	38	0	120
make—do, fix	1074	630	456	900	665	260	600
me, I—I	54	126	0	80	19	20	56

TABLE 16*a*—(*Continued*)

	Weighted frequency of each error						
	Filipino	Japanese	Chinese	Korean	Portuguese	Hawaiian	Kindergarten
more big	177[H]	168	72	100	494	120	240
more bigger	88	126	72	60	494	100	80
no—not	2543	2016	2352	2240	2052	2060	840
no more	1112	882	264	960	684	560	320
no need	56[P]	126	72	100	114	60	120
one—a, an	227[H]	168	384	340	627	240	160
open—turn on, *unfasten*	0	42	48	80	0	160	0
plenty—many	195	252	288	400	380	300	280
scare—be afraid	114	252	192	100	95	20	0
shame—bashful, be ashamed	88	84	96	0	0	80	0
sore—hurt	54	84	72	0	38	220	160
stay—is	752	126	216	420	705	180	400
"take it go"	63	0	72	40	19	40	0
tell, say ask	418[P]	0	144	420	285	40	80
the—how	120	168	72	100	133	20	240
too—very	22	84	0	40	133	80	0
try as auxiliary	342	462	360	360	304	40	160
went go	547	126	168	980	342	180	320
waste time—don't want to	237[H]	588	168	360	114	260	280
wild—scold, angry	94[P]	42	0	0	171	100	0

H error at least twice as frequent for Honolulu children.
P error at least twice as frequent for plantation children.
†Average of Honolulu children x19 and plantation x25.

uses of English words, which, athough not exactly incorrect, occur very rarely in standard English. These words and phrases are listed in Table 16, and sentences illustrating their use are found in Appendix *D*. Since there would be no occasion to use many of these particular phrases during the limited period of observation, the balance would be much upset by a single child's repeated use of a word; so comparisons in this table are made according to the number of children who made the specific error. Then, since the groups differed so much as to amount of English used, the numbers making each error were multiplied by a constant that would give each group equal weight on the basis of amount of English words used. These constants were the smallest two-digit whole numbers that would give approximate equality to the products when used to multiply the number of English words used by each group.

The most common of all these specific errors was the use of *"no"* for *"not,"* which was made by 722 of the 975 children whose records were examined for this table. The probable reasons for this high frequency were discussed above. The only marked difference among the groups as to frequency is for the kindergarten children, only a fifth of whom make this particular error.

"No more," another negative error listed, is rarely used correctly but rather simply to express absence or lack of some person or object.

"Like" for *"want"* is used by almost half the children. It might be considered also as used for *"would like to,"* and the sentences where it occurred were so corrected. The Japanese and Chinese use it most often.

Next in frequency is the use of *"go."* It might be considered simply as a redundancy or as an auxiliary. It is most frequently used to express future time as may be done correctly in English in the phrase, *"I am going to,"* but it is used in present form without either the auxiliary or the infinitive sign. It seems to be used sometimes for emphasis and occurs with the present when any other tense form might be correct.

"Get" is used erroneously by about two-fifths of the children. The Japanese use it least often, probably simply because they use fewer English verbs. It occurs almost entirely in its present form and with considerable extension of meaning, including that of

possession, existence, and occurrence. Colloquial English, of course, uses the word in the sense of possession in the phrase *"I've got,"* but this particular use never occurred among the Island children studied. The overuse of *"get"* is probably due to its early introduction as an easy and important verb into pidgin English and to the meagre vocabulary of our subjects.

"Make," another greatly overused verb, is used wrongly most often by the Filipinos and Koreans. It has had its meaning extended to cover many varieties of action, partly due to limited vocabulary, and is almost always used intransitively.

"Kind" occurs very frequently where it would appear that the child is at a loss for a word to express himself more adequately. It and *"funny kind"* are especially frequent among the Chinese and Japanese. They use it frequently in the place where Chinese would employ a classifier.

Five more words are used wrongly by over 10 per cent of the children.

"Stay," which Reinecke (26) identifies as the Portuguese *"esta,"* is used wherever *"is"* might be in standard English. The Portuguese and Filipino children lead in frequency of its use.

"Try," used before an imperative, is used a little more often by the Japanese. There is such a use of a word that may be translated *"try"* in Japanese, but in that language it follows the verb.

"Plenty" is used for the adjective *"many"* slightly more often by the children of the earlier immigrants than by those of the later.

"Waste time" has an entirely different meaning for the Island group than it does in standard English. It usually means little more than *"I don't want to."* The Japanese use it most.

"Tell," with its meaning extended to include *"ask"* and *"say,"* is especially popular with the Koreans and Filipinos. In Portuguese, the same word is used for *"tell"* and *"say."*

"One" for *"a"* or *"an"* is a natural error to be passed on to the children by adults in whose language there is no indefinite article, or, if there is, it is the same word as for the numeral *"one."* It is not always really wrong, as when the child says, *"I see one man,"* but it sounds strange for in standard English, *"I see a man"* would be much more usual. It is listed here more frequently than in the table of errors, where it was counted only when obviously incorrect.

The Portuguese make this error especially often. It might be considered, therefore, that it is a more difficult error to eradicate if the home language has an indefinite article, but the word is used also for the numeral *"one"* than when the article is completely lacking, were it not for the fact that the Chinese and Koreans who rank next as to frequency have no articles at all in their languages.

"Went" preceding another verb is a favorite method of expressing past tense for the Koreans and is used to some extent by all groups. The Filipinos use *"go," "already,"* and occasionally *"been"* for that purpose, whereas the Koreans besides *"went"* use only *"go."* The Chinese prefer *"already"* after the verb rather than *"went"* for this purpose, perhaps because *"already"* corresponds in its correct use more nearly to the Chinese use of particles after verbs to express completed action. The Portuguese make some use of all methods for expressing past time, using *"went"* and *"go,"* however, more often than *"been"* and *"already."* Few of the Hawaiians use any of these methods; the highest number, nine, use *"went."* The Japanese prefer the use of *"already"* to any other method except *"go."*

No other error was made by as many as 70 children, but among them are certain that show marked differences between the groups. *"That's why,"* more commonly reduced to *"as why,"* is used instead of *"because"* in answering the question *"why"* by over twice as many Filipino, Portuguese, and Korean children than of any other group.

A group of phrases, *"bring it come," "find,"* for *"hunt," "take it go,"* and *"long time,"* are used almost exclusively by the Chinese. The first three are translations of Chinese idioms, although in the third phrase, *"big"* would be a more exact translation than *"long."* As for the fourth, the same word may be used in Chinese for both *"find"* and *"hunt."* The Chinese are second in the use of *"lend"* for *"let,"* another instance where the same Chinese word could be translated either way. This error is, however, made by several children from other groups and even more often by the Japanese. This is probably due to the similarity in sound between the two words. The use of *"scare"* for *"afraid of"* is found mainly among the Japanese and Chinese.

The Japanese are almost the only ones to use *"me, I"* together as the subject of a sentence. It is good Japanese. So is *"house"* in the sense of *"family"* or *"we,"* an error peculiar to that group. *"Broke"* for *"break"* and *"little more"* for *"almost"* or *"soon"* were other errors used by more Japanese children than by the others.

The Portuguese are responsible for almost half the errors in comparison. Colloquial Portuguese occasionally allows such double comparisons as *"more better"* (9, p. 201). The other error, *"more big,"* is made more often by the non-Portuguese—the other languages in which adjectives are compared doing so by means of words similar to *"more." "All same," "chance"* for *"turn," "too"* for *"very,"* and *"wild"* for *"angry"* or *"scold"* are also used by more Portuguese children than by the remaining groups.

The Hawaiian word *"hemo"* could be correctly employed for all the unusual uses of *"open"* which the Island children employ. Many Hawaiian words may be used as both verbs and adjectives, so the use of *"sore"* for *"hurt,"* and of *"lazy"* for *"tired of"*, most frequent among the Hawaiians, would be not unnatural errors for the Hawaiian children.

More Koreans than any of the other groups use *"break"* for *"tear."*

The children of the latest immigrants, the Filipinos, lead clearly in no particular error type. They are too much influenced by the other racial groups around them. They do, however, lead in vulgar, slang, and swear words.

C. SENTENCE LENGTH

1. *Average Number of Words Per Sentence Irrespective of Language Used*

One of the best criteria of language development is the length of the sentence used. We have attempted to make use of this criterion as a measure not only of the portion of the child's speech that was in English, but also of that part that was spoken in other languages.

Miss Chun (8) in her thesis compared English and Chinese sentences expressing the same thought as to length as measured by the number of words in each. She found very little difference; if any, the same thought expressed in Chinese words required more words. To determine if the same would be true of the other languages our subjects spoke, 50 brief childish sentences spoken in English were selected and translated by a representative of each linguistic background from among the workers on the project, in consultation with authorities in that language. The translations were all in simple form as a child would speak to children (see Table 17). The average number of words per sentence were as follows: in English, 3.10; in Ilocano (the Filipino dialect used in translation), 3.14; in Portuguese, 3.63; in Japanese, 4.98; in Hawaiian, 4.16; in Korean, 2.51; and in Chinese, 3.52.

TABLE 17

COMPARISON OF THE AVERAGE LENGTH OF SENTENCES WITH DEGREE OF BILINGUALITY OF THE GROUPS STUDIED

	Japanese	Chinese	Filipino Rural	Filipino City	Korean	Hawaiian	Portuguese
Number of children using sentences entirely in a non-English language	117	84	95	51	50	54	9
Per cent of English words used	50	80	79	92	96	97	99
Average sentence length in all languages	2.7	3.2	3.0	3.4	3.1	3.1	3.3
Average length of the 50-sentence sample in each language	4.98	3.52	3.14	3.14	2.51	4.16	3.63

If there were no difference in progress in speech between the groups studied, then the average length of sentences of each group should, if they used much of languages other than English, place them in the same order as that found to be true for the length of sample sentences translated into their respective languages. Then those using sentences longer than children speaking English only should be the Japanese, Hawaiian, Portuguese, and Chinese, and those using sentences shorter, the Filipino and Korean. But this is not the order found, as can be seen by reference to Tables 18 and 19. The English-speaking Caucasian monoglots rank first; then in order come city Filipino, Portuguese, Chinese, Korean, Hawaiian, rural Filipino, and last of all, Japanese.

Taking as measures of degree of bilingualism of the group the number of children using sentences entirely in a language other than (all children composed at least a few sentences entirely in) and the proportion of English used, it is found that length ence bears a closer and a negative relationship to these s. The Japanese who use English and Japanese in about mounts average sentences of only 2.7 words, even, although expression of thought in their parental language requires the number of words; the two groups using about four-fifths one-fifth another language, use sentences averaging 3.1 the three groups of whom only two-fifths of the children ences entirely in a non-English language and who use 92 er cent English, average sentences 3.2 words in length; and tuguese, of whom only nine children knew enough Portu- frame even one-word sentences in that language, use sen- hat average 3.3 words.

each group is considered as a whole, the average number s per sentence is in every case less than that of the mono- English-speaking children, although very few six-year-olds cluded in that study. This retardation is not marked until rs of age, however. In fact, four out of the seven groups years average higher than the Caucasian monoglots at that is is probably due to the fact that the youngest children are y selected group; for, as explained above, in several groups thers refused to have their babies, who talked very little yet, . All groups who use more than 10 per cent of a non- language are below the monoglot mean at two years. At rs old, only the Portuguese exceed the Caucasian monoglots ce length, while the Chinese and city Filipinos tie with the ns and the rest of the groups all use shorter sentences. The three-year-olds, who average 37.3 months in age, have a ree-year-old mean sentence length than any other group.

very group, the increase in sentence length from two to rs is statistically significant, but after that, the gain is more and much slower. Five of the seven groups make a gain ee to five years that is statistically significant, but of these, Japanese and Hawaiian, lose the next year, and only the two groups continue to gain. The city Filipinos are the only hose gain from five to six years approaches significance.

posite curve of the average of the seven means is 0.1 higher

TABLE 18

COMPARISON OF LENGTH OF SENTENCES USED BY DIFFERENT GROUPS

	At									Differences			
	2	3	4	5	6	All	Boys	Girls	Girl-boy	6-2	3-2	5-3	6-5 yrs.
Japanese													
S.L.	1.6	2.5	2.9	3.3	3.1	2.7	2.8	2.6					
σ dist.	.51	.54	.66	.67	.66	.84	.91	.75					
σ av.	.10	.11	.13	.13	.13	.075	.11	.10					
Diff.									—0.2	1.5	0.9	0.8	—0.2
σ diff.									.15	.17	.15	.17	.18
D ÷ σ									1.3	8.8	6.0	4.7	1.1
Chinese													
S.L.	1.6	3.2	3.4	3.8	3.8	3.2	3.2	3.2					
σ dist.	.54	.63	.34	.65	.75	.98	.92	1.06					
σ av.	.11	.13	.07	.13	.15	.09	.11	.14					
Diff.									0.0	2.2	1.6	0.6	0.0
σ diff.									.18	.19	.17	.18	.20
D ÷ σ									——	11.6	9.4	3.3	——
Korean													
S.L.	1.9	3.1	3.4	3.4	3.8	3.1	3.1	3.2					
σ dist.	1.02	.85	.42	.58	.55	.85	.78	1.30					
σ av.	.20	.17	.08	.12	.11	.076	.10	.16					
Diff.									0.1	1.9	1.2	0.3	0.4
σ diff.									.19	.23	.26	.21	.16
D ÷ σ									0.5	8.3	4.6	1.4	2.5
Filipino													
rural													
S.L.	1.6	2.8	3.2	3.6	3.7	3.0	2.9	3.0					
σ dist.	.59	.54	.57	.61	.58	.97	.98	.95					
σ av.	.12	.11	.11	.12	.12	.087	.12	.12					
Diff.									0.1	2.1	1.2	0.8	0.1
σ diff.									.17	.16	.16	.16	.17
D ÷ σ									0.6	13.1	7.5	5.0	0.6
city													
S.L.	2.0	3.2	3.7	3.9	4.3	3.4	3.3	3.5					
σ dist.	.59	.52	.65	.51	.49	1.05	.92	1.01					
σ av.	.12	.10	.13	.10	.10	.094	.12	.13					
Diff.									0.2	2.3	1.2	0.7	0.4
σ diff.									.18	.16	.16	.14	.14
D ÷ σ									1.1	14.4	7.5	5.0	2.9

TABLE 18—(*Continued*)

	At									Differences			
	2	3	4	5	6	All	Boys	Girls	Girl-boy	6-2	3-2	5-3	6-5 yrs.
Hawaiian													
S.L.	2.1	2.8	3.6	3.8	3.4	3.1	3.0	3.2					
σ dist.	.92	.73	.96	.61	.73	1.05	.91	1.05					
σ av.	.18	.15	.19	.12	.15	.09	.12	.13					
Diff.									0.2	1.3	0.7	1.0	—0.4
σ diff.									.18	.23	.23	.19	.19
D ÷ σ									1.1	5.7	3.0	5.3	2.1
Portuguese													
S.L.	2.5	3.4	3.4	3.6	3.7	3.3	3.2	3.4					
σ dist.	.49	.70	.59	.68	.52	.74	.67	.78					
σ av.	.10	.14	.12	.14	.10	.066	.08	.10					
Diff.									0.2	1.2	0.9	0.2	0.1
σ diff.									.13	.14	.17	.20	.17
D ÷ σ									1.5	8.6	5.3	1.0	0.6
Caucasian													
S.L.	1.8	3.2	4.3	4.7		3.6							
σ av.	.15	.13	.09	.09		.09							
Composite	1.9	3.0	3.4	3.6	3.7								

TABLE 19

DIFFERENCES OF LENGTH OF SENTENCES USED BY DIFFERENT GROUPS

			Difference	σ diff.	D ÷ σ
Monoglot		City Filipino	0.2	.13	1.5
“	—	Portuguese	0.3	.11	2.7
“	—	Chinese	0.4	.13	3.1
“	—	Korean	0.5	.12	4.2
“	—	Hawaiian	0.5	.13	3.8
“	—	*Rural Filipino*	0.6	.13	4.6
“	—	Japanese	0.9	.12	7.5
City Filipino	—	Portuguese	0.1	.11	0.9
“ “	—	Chinese	0.2	.12	1.7
“ “	—	Korean	0.3	.12	2.5
“ “	—	Hawaiian	0.3	.13	2.3
“ “	—	Rural Filipino	0.4	.13	3.1
“ “	—	Japanese	0.7	.12	5.8
Portuguese	—	Chinese	0.1	.11	0.9
“	—	Korean	0.2	.11	1.8
“	—	Hawaiian	0.2	.11	1.8
“	—	Rural Filipino	0.3	.11	2.7
“	—	Japanese	0.6	.11	5.5
Chinese	—	Korean	0.1	.13	0.8
“	—	Hawaiian	0.1	1.2	0.8
“	—	*Rural Filipino*	0.2	.13	1.5
“	—	Japanese	0.5	.12	4.2
Hawaiian	—	Rural Filipino	0.1	.13	0.8
“	—	Japanese	0.4	.12	3.3
Korean	—	Rural Filipino	0.1	.12	0.8
“	—	Japanese	0.4	.11	3.6
Rural Filipino	—	Japanese	0.3	.12	2.5

at two years, 0.2 lower at three years, 0.9 at four, and 1.1 at five years than the Caucasian curve. By this measure, the preschool population of Hawaii is decidedly retarded in speech.

2. *Comparison as to Length of English, Mixed, and Non-English Sentences*

Table 20 shows the averages of the sentence length for three kinds of sentences for each group: sentences entirely in English, mixed sentences, and sentences entirely in the presumptive home language.

For each group except the city Filipino, the mixed sentences aver-

TABLE 20

AVERAGE LENGTH OF SENTENCES SPOKEN BY EACH GROUP ACCORDING TO LANGUAGE USED IN THE SENTENCE

	At 2	3	4	5	6 yrs.	All	Number of children using each kind of sentence
Japanese children							
Sentences entirely Japanese	1.6	2.1	2.2	2.7	2.5	2.2	117
Sentences entirely English	1.5	1.9	2.3	2.7	2.7	2.2	125
Mixed	2.6	3.5	3.7	4.2	4.2	3.6	116
All sentences	1.6	2.5	2.9	3.3	3.1	2.7	125
Chinese children							
Sentences entirely Chinese	1.5	2.5	2.3	3.1	2.7	2.3	84
Sentences entirely English	1.6	2.9	3.1	3.6	3.8	3.0	125
Mixed	2.9	4.2	4.5	4.7	4.3	4.1	87
All sentences	1.6	3.2	3.4	3.8	3.8	3.2	125
Filipino children (rural)							
Sentences entirely Filipino	1.1	1.6	1.5	1.8	1.7	1.5	95
Sentences entirely English	1.6	2.8	3.2	3.6	3.8	3.0	125
Mixed	2.4	3.5	3.7	4.1	3.9	3.5	117
All sentences	1.6	2.8	3.2	3.6	3.7	3.0	125
Honolulu Filipino children							
Sentences entirely Filipino	1.2	1.6	1.6	1.5	1.5	1.4	51
Sentences entirely English	2.1	3.3	3.7	4.0	4.3	3.5	125
Mixed	2.5	3.4	3.5	4.1	4.2	3.5	117
All sentences	2.0	3.2	3.7	3.9	4.3	3.4	125
Korean children							
Sentences entirely Korean	1.2	1.0	1.3	1.6	1.3	1.3	50
Sentences entirely English	2.1	3.1	3.3	3.4	3.8	3.1	125
Mixed	3.2	3.0	2.8	3.6	4.7	3.4	81
All sentences	1.9	3.1	3.4	3.4	3.8	3.1	125
Hawaiian children							
Sentences entirely Hawaiian	1.1	1.5	1.1	1.2	1.4	1.3	54
Sentences entirely English	2.1	2.8	3.7	3.7	3.5	3.1	125
Mixed	2.6	3.0	3.7	3.7	3.5	3.2	94
All sentences	2.1	2.8	3.6	3.8	3.4	3.1	125
Portuguese children							
Sentences entirely Portuguese	1.0	1.0	1.0	1.5	2.4	1.5	9
Sentences entirely English	2.5	3.4	3.4	3.6	3.7	3.3	125
Mixed	3.0	3.7	3.4	4.2	4.3	3.8	37
All sentences	2.5	3.4	3.4	3.6	3.7	3.3	125

age longest, which is to be expected as no one-word sentences could be included in that group. The Portuguese and rural Filipinos tie; in every other case, the difference between length of mixed sentences and English sentences is greatest for the more bilingual groups.

Only in the case of the Japanese do non-English sentences compare in length with English sentences. For all other groups except the Chinese, the sentences containing no English words average less than one-and-a-half words in length, indicating that many of them are only one-word sentences. Examination discloses the fact that many of these sentences are merely naming some object, calling someone—"*mother*" or a proper noun—or an interjection. So as soon as these children begin to frame longer sentences, they appear to be unable to do so completely in their ancestral language.

The Chinese and Japanese children at two years use sentences of almost identical length in their home language and English; but, whereas this correspondence occurs only at this age for the Chinese children, who thereafter use longer sentences in English, the Japanese children use longer Japanese sentences at three years, very slightly longer English sentences at four, sentences of the same length in both languages at five. At six years, when most of the Japanese group have started to school, their English sentences average 0.2 words longer.

Were it not that the study of the translated sentences had shown that simple, correct speech required as many or more words to express a thought in all languages studied except Filipino and Korean as does English, it might be thought that the briefer sentences when the children were not speaking English were due to the language spoken. As it is, it is more probably due to less facility in the other language.

The sex difference as to sentence length is very slight. It might have been greater had the two-year-old groups been better divided between the sexes, for the boys in every group but one averaged older. However, as in previous studies, what difference there is is in six of the seven comparisons in favor of the girls.

D. ANALYSIS OF SENTENCES ACCORDING TO FUNCTION AND STRUCTURE

Analysis of the sentences used by our subjects included all sentences, no matter in what language or mixture of languages they were spoken. This was possible as all sentences had been translated into English by the recorders and were accompanied by sufficient notes to make the situation clear.

1. *Degree of Egocentricity Shown*

To determine the degree of egocentricity, Fisher's (11) method was used. A sentence was classed in Category I if the speaker was the subject; in Category II if any other person was; in Category III if an animal or thing was; and as non-verbal if it did not make an intelligible use of any of the six languages used by our subjects. As some of the languages spoken are so much more elliptical than English, the sentences were classified under different categories according to subject understood rather than subject named. When no subject was clear, the sentence was classed under Category III,

TABLE 21

INDICES OF EGOCENTRICITY FOUND FOR EACH GROUP

				Filipino			
	Chinese	Japanese	Korean	City	Rural	Hawaiian	Portuguese
At age							
2	30	43	36	19	14	37	22
3	46	43	43	45	43	42	44
4	49	42	39	44	44	63	48
5	46	52	41	36	41	66	41
6	44	45	43	40	42	65	28
For							
Boys	40	44	44	40	39	56	37
Girls	47	45	37	34	38	53	37
All	43	45	40	37	38	55	37
Categories							
I	14	13	13	12	13	16	12
II	22	19	23	28	23	25	24
III	10	13	11	9	12	7	13
Non-Verbal	2.6	3.4	2.1	0.9	1.5	1.1	0.3
Yes or no only	1	1	†	†	†	1	†

†Less than one per cent.

as was done by Fisher, except when the sentence was merely "*yes*" or "*no*." The results of this analysis are shown in Table 21. The individual variability for all categories and for the egocentric indices, which are obtained by dividing Category I by the sum of Categories II and III, is very large.

Our results agree with Fisher's on monolingual Americans in that in six groups, Japanese being the exception, the average index is lowest at two years. The Japanese sentences were the most difficult to classify on the basis of subject, as correct Japanese does not necessarily involve the naming of the subject in any type of sentence. It was necessary if the results from analysis of the different languages were to be uniform that the sentences be classified under the different categories according to subject understood rather than according to subject named. It is possible that this was more easily done at two years than later for the Japanese sentences. In four groups the average index drops from five to six years, and in two more it is lower at both five and six years than at four years, although the average at six years is greater in both cases than at five. The exception in this case is the Korean group.

The average indices of egocentricity and number of sentences in Category I for each group as a whole show considerable similarity. The Filipino and Portuguese groups have almost identical averages for both these measures. The three Asiatic groups have very similar averages for these two measures, the slightly higher Japanese egocentric index being due to the unusually high average of the two-year-olds. The largest difference in number of sentences in Category I for the first three groups named above is only 0.7; for the last three groups it is 0.4.

The Hawaiian show the largest number of sentences with the first person as a subject and rank second as to number of sentences in Category II or those sentences where another person is the subject. They are the group which showed most interest in the examiner and turned away most often from their play to enter into conversation with her. It may be that there is a real difference in the degree of interest in persons indicated here.

The two Filipino groups show little similarity in the proportions of sentences in Categories II and III. The country group is more interested in things and animals and the city group in persons, if the subjects of sentences can be considered a measure of interests.

Japanese and Portuguese use the largest number of sentences in Category III, the city Filipinos and Hawaiians the fewest. Indeterminate sentences are, according to Fisher's rules, thrown in that category.

In only one group does the number of sentences classified in Category III fail more than once to decrease with rise in age level. Sentences classified under Category II tend to increase with age but not so uniformly, while after two years, those in Category I remain fairly constant in half of the groups, dropping off from five to six years.

Most of the sentences classified as non-verbal are found at two years. Their number seems to bear a relation to degree of bilingualism. The three Asiatic groups, in most of whose homes two languages are heard and two of which use the least English at two years, use the most non-verbal sentences. If city Filipinos and Hawaiians were to change places in regard to frequency of non-verbal expressions, the resulting order of rank would agree exactly with those for amount of English used at two years and also with language ratings of homes. It would appear, therefore, that there is a greater tendency for bilingual children to resort to such expressions in lieu of true words than there is for monoglots.

2. *Classification According to Grammatical Types*

When the sentences are classified according to the four grammatical types (see Table 22), there is a slight increase with age in interrogative sentences and a decrease up to four years in exclamatory sentences. This is exactly what was found in our previous study of monolingual children (31), but the proportion of questions is slightly less, that of exclamatory sentences more. Declarative sentences increased slightly until five years, imperatives until four years; but if sentences not imperative in form but expressive of wishes, requests, and commands are counted also, as is done under the heading "Variations of Imperative," the number is greatest at two years old.

There is much less variability among the racial groups if variations of imperatives are included than if they are omitted. The per cent is then 28 or 29 (it was 28 for our monoglots) for five groups, the Chinese children using considerably less and the city Filipino more than the average. Criticism and threats were not tabulated separately for all groups. They were found in three per cent of the sentences for each group for which they were tabulated.

TABLE 22

FREQUENCY OF VARIOUS TYPES OF SENTENCES USED BY THE SEVEN MAIN GROUPS STUDIED

	Chinese	Japanese	Korean	Filipino		Hawaiian	Portuguese	At					Boys	Girls
				City	Rural			2	3	4	5	6		
Declarative	59	60	61	58	63	59	65	56	61	62	64	62	59.9	62.2
Interrogative	10	10	11	9	8	16	8	6	10	11	12	13	10.2	10.6
Exclamatory	12	9	7	4	5	6	2	13	6	4	4	4	7.1	7.2
Imperative	17	20	20	28	24	19	24	20	23	26	21	22	22.3	20.2
Imperative and variations of imperative	23	29	29	34	29	29	28	32	29	30	26	26	29.8	27.8
Answers	6	9	7	3	5	11	6	5.2	6.2	6.7	7.7	7.3	6.5	6.5
Naming only	4	4	1	3	3	4	3	9.3	3.6	2.5	1.3	1.2	3.1	3.4
Criticism and threats			3	3	3		3							
Negative	12	12	12	13	13	10	13	9.7	13.1	12.1	13.0	13.1	12.0	12.6
Complex and compound	4.4	5.1	5.1	7.5	5.2	4.0	3.6	0.9	3.6	4.4	7.3	7.5	4.4	5.0
Incomprehensible	2	0.6	2	0.7	1	1	1	4.7	0.6	0.2	0	0	0.67	0.46

For every group, questions are more numerous than answers, perhaps because questions were frequently addressed to the adults present who were refraining from questioning the children during the period of observation. Answers, like questions, tend to increase in numbers with increase in age, as the child's conversation becomes more truly an interchange of ideas with another person.

Merely naming objects or persons, which composes over nine per cent of the sentences at two years, falls to 1.2 per cent at six years. The percentage is not quite so high as that found for our monoglots —13.2 per cent at two years—but it decreases more gradually, not reaching the level found for them at four years until a year later.

Negative sentences are more numerous after two years. The average per cent for each racial group is 12 or 13 (for our monoglot series it was 12) except for the Hawaiian. They used fewer, as was the case in Situation *A* in the previous study. As mentioned above, the Hawaiian children paid much attention to the observer, and so the situation for them was more nearly comparable to Situation *A* than it was for the other groups. This probably also explains the higher per cent of interrogative sentences used by the Hawaiians.

If the use of negative sentences and clear-cut imperatives be considered indicative of more aggressive and dominant personalities, the combined per cents of these types of sentences might be used as such a measure. According to this criterion, the Filipinos and Portuguese would be the most aggressive, the Chinese and Hawaiians least, and the Japanese and Koreans would fall in between. Curiously enough, the order of conviction of the different races in the Territory for such crimes of violence as manslaughter and murder is from most to least Hawaiian, Filipino, Korean, Japanese, and Chinese, when measured by the rate of convictions to each thousand males 18 years of age or over (2, p. 34). This order, with the exception of the Hawaiians, is exactly that for the ranking of our subjects as to aggressiveness when measured by the sum of negative and clear-cut imperative sentences.

The children whose homes have higher language ratings and who use more English tend to use fewer exclamatory sentences, but the correlation is not at all close.

The proportion of complex and compound sentences, which was calculated for English sentences only, rises with increase of age; but at no age nor for any group is it as large as for the monoglots, whose use of such sentences rose from 1.6 per cent at two years to 17.1 per cent at six years. The two highest averages are found for two of our most monolingual groups. The third such group, the Hawaiians, however, has the lowest average. That may be in part due to their excessive proportion of answers which are usually simple sentences and often single words.

E. SPECIAL ANALYSIS OF QUESTIONS

1. *Comparative Frequency of Questions*

A special study was made of all questions asked by the children, using the modification of Isaac's (15 Appendix) classification employed in a previous study (29). The results of this analysis are shown in Tables 23 and 24, with that of the questions asked by the Caucasian group previously studied shown for comparison. The present analysis is based on 5,092 questions and includes all questions asked by each child, irrespective of language used.

Although questions increase in proportion to sentences used with age and the other Caucasian group included very few children at the six-year-old level, only one group, the Hawaiian, asked as many questions per child as did the Caucasians in Situation *C*.

A higher proportion of children asked questions in the Island groups, probably due to the fact that the samples of the Caucasians had been taken on a time basis and some used very few sentences. In per cent of sentences used that were questions, the other Caucasian and the Hawaiian children surpassed all the other Island groups. The Hawaiian children not only exceeded the other Caucasians in Situation *C* but almost equalled those in Situation *A* as to proportion of questions; however, it must be remembered that they were a little older and they directed 73 per cent of their questions to adults, whereas the other Caucasians in Situation *C* asked adults only 28 per cent of their questions. Probably the greater frequency of questions among the Hawaiians is due to their extreme interest in the observer, whom they addressed very frequently.

Only the rural Filipinos failed to ask more questions of adults than did the Caucasians in Situation *C*; and this was in spite of the fact that the Island groups are slightly older and that the tendency to turn to adults when playing with other children decreases with age. The average number of questions asked of adults by all groups at six years is very nearly, but still slightly higher than, the average of all Caucasian children in Situation *C*. This tendency may be due to the smaller groups of children observed.

The Korean and Portuguese ask fewer questions seeking corroboration and approbation, such as, *"I go use all, all right?"*, *"us two pals, eh?"*, than did the other Caucasians. The city Filipino asked more and the rural Filipino less, so that the average of these two groups gives, as do the rest of the seven main groups, approximately the same proportion of such questions as were asked by the Caucasians.

A comparison of all the boys with all the girls from the main groups showed similar sex differences to those found in the previous study. The boys asked more questions and a higher per cent of causal questions (those of reality other than place) and fewer of action, as was true in the previous study; but whereas the Caucasian girls had asked more questions of place, the per cent of such questions was almost identical in this study for the two sexes; and the girls, not the boys, asked the greater number of questions of calculation.

2. *Analysis According to Function*

When all seven groups are combined, certain age trends as to question type are discernible that are almost identical with those found for the Caucasian monoglots. There is considerable irregularity in these trends, however, when each group is taken separately, except in the case of the Portuguese who show trends very similar to the other Caucasians.

Four types of questions show an increase in proportion with age. They are causal, illustrations of which are: *"Why I have to bathe?" "Why you no like play?" "Why you broke 'em?" "This for what?"*; regarding rules, *"How you pray?"*; in calculation, *"How many spoon, Nang* (Mother)?"; *"Fifth, eh?"*; and other actions, *"Who like box me?"*; *"You going D's house?"*; *"She never come home, yeah?"*; and *"She no make fall down, no?"*

Three types of questions show a decrease with age. They are regarding names of objects, such as, *"What this is?"*; *"What that?"*; regarding place, *"Where da knife?"*; *"Where get?"*; *"Where stay broke?"*; and merely asking for repetition of remark, as *"Ha?"*; *"What you said?"* This last type is referred to as the *"what say"* type.

Two types increase in frequency to four years and then decrease.

TABLE 23

COMPARATIVE SUMMARY OF QUESTIONS USED BY EACH RACIAL GROUP

	Per cent of each type classified according to form and function											
	Other Caucasians Situation		Chinese	Japanese	Korean	Filipino		Hawaiian	Portuguese	Situation A	Rural Japanese	Kindergarten
	A	C				City	Rural					
I Causal	11	5	5	2	4	3	5	4	6	2	1.5	7
II Reality												
a. Place	10	15	18	15	17	15	18	13	17	22	27	11
b. Others	17	12	11	13	11	12	12	14	18	13	26	14
III Actions												
a. Imperatives	6	14	5	5	3	5	4	8	3	8	5	6
b. What say	4	7	14	14	9	4	10	6	5	4	11	4
c. Others	34	28	35	31	40	46	37	43	39	36	18	40
IV Name	5	6	3	8	6	8	6	4	6	8	1.5	3
V Classification and Evaluation	9	11	7	11	7	8	6	6	4	6	11	14
VI Rules	1	2	0.5	0.5	0.1	0	0	0.1	0	0	0	1
VII Calculation	4	1	1	0.6	1	0.3	1.2	0.6	2	0.8	0	0.5
Corroboration and Approbation	23	14	13	14	5	17	9	13	8	6	24	16
Asked of adults	100	28	33	29	51	38	25	75	42	100	—	87

TABLE 23 (*continued*)

	Other Caucasians Situation A	Other Caucasians Situation C	Chinese	Japanese	Korean	Filipino City	Filipino Rural	Hawaiian	Portuguese	Situation A	Rural Japanese	Kindergarten
Only English questions considered												
Interrogative words	41	39	53	39	45	40	54	32	49	46	73	42
What	14	16	19	16	16	18	20	7	16	14	27	13
Where	7	11	17	9	16	14	17	12	14	22	18	12
Who, which, whose	2	3	7	8	5	5	9	3	7	4	18	6
How	9	3	5	3	2	1	2	2	3	6	8	4
When	2	0.6	0.2	0	1	0	0	4	4	0	0	0.3
What for	2	1	1	0	0.3	0.2	0	2	1	0	0	3
Why	5	2	3	2	4	2	6	1	4	1	0	4
With yeah, no, you know, eh, etc.	—	—	10	28	6	17	13	5	6	—	—	12
With Japanese or Chinese† enclitic	—	—	4	2	—	—	—	—	—	—	—	4
Total No. of questions	1529	1556	633	640	694	573	479	987	528	120	66	372
Av. No. per child	10.4	7.9	5.1	5.1	5.6	4.6	3.8	7.9	4.5	4.4	3.9	4.5
% of children asking questions	88	78	90	96	96	95	86	99	93	92	94	87
% of sentences that were questions	16.2	11.6	10.1	10.2	11.1	9.2	7.7	15.8	8.5	9.6	7.9	9.1

†% of English and mixed sentences

TABLE 24

COMPARATIVE SUMMARY OF QUESTIONS CLASSIFIED ACCORDING TO FORM AND FUNCTION AS USED AT DIFFERENT AGES

	Average at each age level and by sexes						
At	2	3	4	5	6 years	Boys	Girls
I Causal	0.4	3.5	3.5	4.5	6.7	4.4	3.5
II Reality							
a. Place	20.0	18.3	16.0	14.4	14.5	16.1	15.9
b. Others	9.3	10.6	14.7	14.1	13.3	14.6	12.3
III Actions							
a. Imperatives	3.3	5.0	6.1	5.3	3.1	4.1	5.3
b. What say	20.1	9.9	7.6	6.7	6.4	9.0	8.6
c. Others	30.3	36.1	38.4	43.1	41.4	37.7	39.9
IV Name	10.6	8.7	4.9	4.4	5.1	6.3	6.1
V Classification and Evaluation	4.9	7.7	7.4	6.0	6.9	6.7	7.0
VI Rules	0	0.1	0	0.4	0.8	0.2	0.5
VII Calculation	0.2	0	1.2	0.8	2.1	0.8	1.1
Corroboration and Approbation	10.9	11.6	9.6	11.0	12.1	10.6	11.8
Asked of adults	72	49	44	32	30	41	43
Interrogative words	36	43	45	46	48	46	44
What	13.0	17.1	17.7	15.1	15.7	15.6	16.4
Where	18.1	15.7	13.1	12.7	12.7	14.6	14.0
Who, which, whose	2.8	4.4	6.3	9.1	7.8	7.2	6.1
How	1.5	1.5	2.0	3.0	4.4	2.5	3.0
When	0.8	1.0	2.1	1.4	1.1	1.3	1.1
What for	0.1	0.4	0.9	0.6	0.9	0.9	0.4
Why	0.1	2.4	2.5	3.9	5.0	4.0	2.7
With { you know yeah, no, eh etc. }	6.4	10.0	9.4	8.4	10.3	9.0	10.0
Total no. of questions	538	882	960	1045	1109	2286	2252
Per cent of children asking questions	69	95	98	97	98	94.1	93.4
Per cent of sentences that were questions	6.1	10.0	10.0	11.9	12.7	10.2	10.6

They are imperative, *"May I have one?"*; and regarding reality other than place. The latter includes several sub-types: time, *"George Washington born tomorrow?"*; *"When he went born?"*; *"Inang* (mother) *what time pau?"*; facts and events, *"As* (is that) *all the food?"*; *"Mama no more one penny?"*; *"Whose chance for deal* (turn to deal) ?"; *"Tomorrow get* (will there be) *church?"*

Questions of classification and evaluation rise in proportion from two to three years, after which there is decrease to five and then again an increase. Such questions are, *"You Korean, no?"*; *"Dry kind, eh?"*; *"My fingernail big, no?"*; *"Dis asin* (salt)?" But the more mature types of questions are rarer than they are with the Caucasians. Questions regarding number and rules are almost nil until four and five years, and even at six they are very infrequent. As to causal questions, the average for all is less at five years than for the Caucasians, and only the Portuguese group as a whole exceed the Caucasians, due to a high average (nine per cent) of such questions at six years. Of the question types that, for the Caucasian children, increased with age up to five years, only on those dealing with other actions do the Island groups excel; on the other three types they average lower.

On the three question types that decrease in proportion as age increases, most groups average higher than do the Caucasians. The four most bilingual groups ask to have statements repeated much more often than do the Caucasians and the more nearly monolingual Island groups. This is probably due to greater difficulty in understanding when they are learning two languages. Only the Chinese and Hawaiian, a nearly monoglot group, fail despite their advantage in age to equal or exceed the Caucasians in asking for the names of things. Only the Hawaiians ask fewer questions of place.

In general, so far as question types go, the preschool children in Hawaii show less maturity than did the Caucasians previously studied, and the more bilingual groups show more often a need to inquire as to names of objects or to have statements repeated to them. All groups except the monolingual Portuguese show less maturity in the use of questions of causality.

3. *Analysis According to Form*

Questions of corroboration in their form show the effect of pidgin English and maternal languages. The Chinese and Japanese children make considerable use of their home languages' method of turning declarative sentences into interrogative by the addition of a syllable that is commonly used for that purpose. Sometimes such particles are attached to sentences otherwise entirely in English. All groups make use of *"yeah," "no"* (for "isn't it so?", which is also very similar to the Japanese interrogative), *"eh," "ha,"* (also a Chinese interrogative), and the Japanese and Honolulu Filipinos use *"you know"* to turn declarative sentences into interrogative. Not one makes use of *"isn't it"* or *"don't you"* and similar standard English phrases when seeking approval or corroboration of remarks.

The use of English interrogative words when asking English questions was also studied. In the previous study, the use of such words was found to decline with age.

In these groups, the use of such words increases slowly but steadily with age, at five years being used in exactly the same proportion as by the Caucasians at three years, and at six years in 48 per cent of English sentences as against 49 per cent for two-year-old Caucasians. *"What"* and *"who"* are more often used by each group except the Hawaiian, and *"where"* except by the Japanese, than these words were by the Caucasians. Although so many more interrogative words are used by the polyglot groups, in the case of those interrogatives previously found to increase with age and which (except for *"when"*) do in this study also, the proportions of *"how"* questions are larger only for the Chinese, of *"when"* only for the two most nearly monoglot groups, and *"what for"* only for the Hawaiian. *"Why"* is used more by most groups, even though their causal questions are fewer.

F. ANALYSIS OF CONVERSATIONS BY PARTS OF SPEECH

1. *Parts of Speech Used in English*

The conversations were next analyzed, as had been done in previous studies, as to parts of speech. In Table 25 are shown the results of analysis into the different parts of speech of the English words as used by the seven main groups studied with the results of the same analysis previously made on the Caucasian group (31).

Under the head of connectives are included conjunctions, prepositions, and relative pronouns, adverbs, and adjectives. As the latter are counted also under the other heads, the sum of the per cents may exceed one hundred. Parts of speech which at some ages fell below one per cent were calculated to tenths, as were the interjections, so that the trends might be clearer. In other cases, the calculation was made only to the nearest round number.

The Caucasian monoglots at each age used approximately 40 per cent of substantives, but whereas the nouns were over twice as many as the pronouns at two years, the proportion of such words decreased

TABLE 25

Comparison of Groups as to the Use of Different Parts of Speech and of Inflection in English

% Eng. words that were	Japanese at 2	3	4	5	6	All	Chinese 2	3	4	5	6	All	Filipino Rural 2	3	4	5	6	All	Filipino City 2	3	4	5	6	All
Nouns	20	20	21	19	17	19	28	20	19	18	15	20	21	15	15	11	14	15	22	18	15	15	15	17
Pronouns	13	24	23	25	24	22	11	17	20	20	23	18	25	22	25	24	24	22	18	25	26	24	25	23
Verbs	17	19	18	22	25	20	25	27	28	26	27	26	14	33	22	33	32	31	26	32	31	32	31	30
Adverbs	18	16	16	14	17	16	17	17	15	16	14	16	25	17	16	18	15	18	16	14	15	14	14	15
Adjectives	12	7	11	10	10	10	6	6	8	8	11	8	4	5	6	6	6	5	4	5	5	6	7	5
Articles	0.1	1.0	1.1	1.4	1.3	1.0	0.6	1.6	2.3	2.3	3.0	1.7	0.4	1.9	2.3	3.6	3.8	2.4	0.8	3.0	2.9	2.5	2.9	2.4
Interjections	19.0	12.3	8.8	7.3	6.1	10.1	13.6	7.8	3.4	4.4	4.6	6.8	9.4	3.5	3.1	2.1	2.5	4.1	12.5	3.8	2.7	3.4	2.1	4.8
Connectives	0.2	0.9	1.2	2.7	2.0	1.4	0.6	3.0	3.6	4.7	3.7	3.4	0.2	2.0	1.9	2.8	3.2	2.0	1.0	2.0	2.3	3.3	3.6	2.4
Copulas	0.2	0.2	0.4	0.4	0.4	0.3	0.6	0.6	0.8	1.0	1.2	0.8	0.6	0.7	1.0	0.7	0.9	0.7	0.4	0.7	0.8	1.2	0.8	0.7
% of verbs *that were*																								
Infinitives	0.1	0.1	0.8	0.4	0.8	0.4	0.0	0.4	1.5	3.3	2.2	1.5	0.0	0.0	0.2	0.4	0.4	0.2	0.0	0.2	0.2	1.0	0.2	0.3
Auxiliaries	3.0	2.6	5.2	3.6	4.9	3.8	0.2	2.7	3.1	4.6	5.4	3.2	0.4	1.8	4.2	3.2	4.9	2.9	0.6	2.7	4.0	4.2	4.2	3.1
Past	6.7	9.4	4.7	6.4	2.1	5.0	7.7	3.1	2.6	2.6	3.8	4.0	4.6	1.9	1.5	2.3	2.3	2.7	3.2	2.9	1.8	2.8	2.7	2.7
Future	0.0	0.0	0.2	0.0	0.1	0.1	0.0	0.0	0.2	0.1	0.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0.2	0.2	0	0.1
Present	0.2	5.0	3.7	6.3	2.4	3.1	0.6	4.6	6.2	6.4	7.1	5.0	1.4	3.2	3.3	4.4	3.7	3.2	1.0	4.6	5.2	5.4	5.9	4.4
Participles Past	0.0	0.1	0.7	0.1	0.2	0.2	0.0	0.0	0.1	1.0	0.8	0.4	0.2	1.0	1.4	0.5	2.3	1.1	0.4	0.4	0.6	0.5	0.4	0.4
Congugated	10.1	11.0	15.3	17.1	11.2	12.9	8.5	10.7	13.8	18.1	19.9	14.2	6.6	7.8	10.7	10.7	13.9	10.0	5.0	10.7	12.1	14.1	13.3	11.0
Per cent of Nouns inflected	2.5	1.8	4.3	2.7	4.6	3.2	0.6	3.4	5.0	6.4	10.7	5.2	3.9	2.6	4.4	6.1	5.4	4.5	0.3	3.4	2.6	4.8	5.0	3.2
Articles definite	75	100	89	94	97	93	100	100	91	91	91	94	94	100	99	99	99	99	89	100	96	97	98	97
Adverbs that were of place or modal	70	82	75	69	68	72	79	73	74	68	62	71	77	67	63	63	63	67	80	73	62	75	73	73

TABLE 25a

COMPARISON OF GROUPS AS TO THE USE OF DIFFERENT PARTS OF SPEECH AND OF INFLECTION IN ENGLISH

(Yrs.)	Korean						Hawaiian						Portuguese						Caucasian				
Per cent of all English words that were	at 2	3	4	5	6	All	2	3	4	5	6	All	2	3	4	5	6	All	2	3	4	5	All
Nouns	21	20	18	15	16	18	34	23	23	23	21	25	28	21	18	19	18	21	28	18	16	15	18
Pronouns	18	22	23	26	24	23	14	21	22	22	22	20	15	23	21	21	22	20	13	23	24	24	22
Verbs	27	31	30	30	29	29	25	27	28	27	26	27	27	25	28	29	27	28	23	26	26	27	26
Adverbs	20	15	13	13	14	15	14	15	14	13	16	14	16	15	16	13	14	15	18	16	12	12	14
Adjectives	6	7	8	7	8	7	4	5	5	6	6	5	7	5	8	9	9	8	4	6	7	8	6
Articles	0.6	2.5	2.7	3.3	2.6	2.4	0.9	2.0	3.1	3.8	3.2	2.6	1.4	2.7	3.2	3.1	3.1	2.7	1	3	4	5	4
Interjections	6.2	2.9	2.0	2.2	2.8	3.2	7.0	3.8	1.4	2.0	2.2	3.3	3.7	1.4	1.7	1.7	1.3	2.0	10.2	3.5	2.1	1.8	4.6
Connectives	1.0	3.4	3.5	3.5	4.7	3.2	1.1	2.7	3.8	4.1	4.1	3.1	1.6	3.4	4.9	3.5	6.7	4.4	2.1	6.2	9.2	9.8	6.6
Copulas	0.8	0.7	1.1	1.4	1.1	1.1	0.2	0.9	1.5	1.5	1.1	1.0	0.5	1.4	0.9	2.0	1.9	1.4	2	3	4	5	3.7
Per cent of verbs that were																							
Infinitives	0.3	0.7	0.6	0.8	1.1	0.7	0.0	0.3	2.1	1.2	1.2	1.0	0.2	1.4	1.4	1.2	1.5	1.1	2	5	7	7	5.3
Auxiliaries	1.0	3.3	4.8	5.7	4.4	3.8	0.5	3.2	6.7	4.0	5.1	3.9	3.4	4.0	6.0	6.3	5.3	5.1	3	8	13	16	9.7
Past	2.7	3.6	4.2	5.2	5.6	4.2	0.8	3.2	2.0	4.8	2.7	2.7	2.8	4.2	4.6	7.6	8.1	5.5	6	5	7	10	7.0
Future	0.0	0.0	0.0	0.4	0.0	0.1	0.0	0.0	0.2	0.0	0.0	0.5	0.1	0.4	0.0	0.3	0.1	0.2	0.2	0.6	1.3	3.0	2.2
Present	3.2	7.1	9.6	12.4	11.4	8.7	3.9	7.7	10.8	11.9	11.1	9.1	4.0	7.0	7.4	11.8	10.7	8.2	3	6	6	6	5.0
Participles Past	0.4	0.2	0.8	0.4	0.8	0.5	0	0.4	1.5	0.9	1.8	0.9	0.7	0.7	0.2	0.5	1.2	0.7	4	2	1	1	1.9
Conjugated	7.7	14.9	19.9	24.7	23.1	18.1	5.2	14.7	23.5	23.1	21.8	17.6	11.2	17.8	19.5	28.0	27.2	20.7	18	27	59	47	32
Per cent of nouns inflected	4.3	3.6	3.4	5.2	6.3	4.1	2.7	3.0	3.6	7.8	5.4	4.4	4.5	8.2	8.2	11.8	10.0	7.8	3.7	4.7	9.3	10.8	7.4
Articles definite	100	99	98	95	99	98	77	92	97	93	98	94	82	92	95	94	97	93	17	37	53	44	42
Adverbs that were of place or modal	77	75	70	47	60	66	87	82	67	68	66	74	82	78	66	54	52	66	86	76	71	66	74

very rapidly the next year and continued to decrease very slowly after that, while the pronouns increased. The proportion of substantives for all the groups in Hawaii approximates two-fifths of all words except that for the Hawaiians the average for all ages is 45 per cent, a little higher. All groups except rural Filipinos show a trend toward a decrease in the proportion of nouns and an increase in that of pronouns with age; but if the per cent of Filipino nouns and pronouns used at each age is averaged with the per cent of English nouns and pronouns weighted according to the amount of each language used at that age, the same trend is found to exist up until six years of age. The average number of nouns then decreases from 29 per cent at two years to 20 at five, while the pronouns increase from 17 to 20 per cent during the same interval.

As the per cent of nouns used in the non-English language is much higher and that of pronouns used much lower in all languages analyzed (see Tables 8 and 26), it would appear that on the whole all the bilingual groups (despite their advantage in age) are retarded in the age at which the shift from nouns to pronouns is completely made. The Japanese, who show one of the larger averages in per cent of English pronouns, would show a much lower average were it not that in very many of their mixed sentences the only English word was a pronoun, usually *"me."*

The per cent of verbs remains nearly constant from three to five for the Caucasian group; the two-year-olds, however, use three per cent less than the average. The same trend is found in all seven groups, except that the lower per cent of verbs continues until four years of age for the Japanese, which is the only group to use fewer verbs than the Caucasian. The latter used 26 per cent of such words, the groups in Hawaii from 26 to 31 per cent. The larger proportion of verbs used by the bilingual groups might be due to the lower proportion of other parts of speech used in English. They comprise a lower per cent of the words used in the other languages.

The Island children, except the Japanese, tend to use a larger proportion of adjectives and a lower proportion of adverbs with increase in age, just as the Caucasians do. Even for the Japanese there is a suggestion of this trend. Most groups use a larger proportion of these modifiers than do the Caucasians, but the difference is not great. The excess of adjectives found for some groups is probably due in part to the very low per cent of articles, for a common error is to substitute the adjective "one" for the indefinite article. There are no articles in Japanese, Chinese, and Korean. The children who speak the first two languages make least use of articles and the Koreans tie for third place from the bottom in frequency of use of these words. None of the other four groups, however, uses at any age as high a proportion of articles as do the Caucasian four-year-olds. For every group, the definite article comprises over 90 per cent of the articles used but less than half as many as used by the Caucasians. So except for the Japanese and Chinese, the discrepancy is due not to failure to use the definite article but rather to the substitution of "*one*" for "*a*" or the omission of the indefinite article. The few times an indefinite article was used by the younger children, it was in a learned phrase.

Connectives are also much less used by the Island children. Only the Portuguese at any age use as high a proportion as the other Caucasians did at three years. A comparatively high proportion of connectives and articles is associated with greater correctness and facility in the use of English [Zyve's (37) third-graders used seven per cent each of articles and connectives] which the groups studied lack.

Interjections, on the contrary, are associated with inability to express one's self adequately in more exact words. Such words are used with greater frequency by the more bilingual groups (except for the rural Filipinos) and the two-year-olds than by the Caucasians of the same age levels. The less bilingual groups either approximate the same proportion or use fewer interjections than do the Caucasians.

2. *Inflection of Words*

Considering the use of verbal forms, we find the Island children to be much retarded. Copulas are rarely used, but only in Portuguese is there a truly comparable word. The Portuguese children use this form most; at five and six years they use it as often as the Caucasians do at two years. Pidgin English ignores the copula. There is little change of the verb to express tense in pidgin English, tense being expressed only by auxiliaries or more frequently ignored. All languages represented are more or less unlike English in their conjugations. Perhaps Portuguese is most like, and we find the Portuguese to be the only group which, whether by auxiliaries or otherwise, conjugates verbs at any age (and then not until six) as much as the three-year-old Caucasians. The Japanese and Filipino groups do not even reach the Caucasian two-year-old norm. The difference is greatest in the use of the sign of the infinitive, auxiliaries, futures, passives (used by only three or four children in all), and past participles. The past tense would also rank very low were it not for the fact that the verb "break" is usually used in pidgin only in its past form, and as it was one of the very few English verbs used by the younger children who did not use a large proportion of English, this one use of the past distorts the picture, particularly in the case of the Japanese. Present participles were used more often by the Island children than by the Caucasians.

Only the Portuguese children inflect nouns to the same degree as the Caucasians, but theirs is the only language represented that forms plurals as does English. Some of the other languages, including pidgin English, have no number for nouns.

The one classification according to meaning is that of adverbs, for which the previous study (30) had found a decreasing proportion of modal and place adverbs as the child became able to use the other types. The per cent of such adverbs for the Caucasians is 74, for the other groups the range is from 66 to 74. In general, the trend is in the same direction as previously found.

Some study was made of the use of comparison of adjectives and adverbs. Except for the use of the word "*more*" usually incorrectly, there were few cases of comparison found, and these almost all were made incorrect by the addition of a superfluous "*more.*"

In general in the use of the different parts of speech and of inflected forms of words, the Island children are retarded.

3. *Partial Study of Parts of Speech Used in Other Languages*

It was desired to discover if any of this apparent retardation was compensated for in the other language spoken. Miss Yee (36) had analyzed the Chinese words into parts of speech for her thesis. The records of the rural Filipinos were similarly analyzed by Mr. Luis, and the results, recalculated to correspond more closely to the English analysis and to each other, are presented in Table 26. It has not yet been possible to complete the corresponding Japanese analysis. None of the other languages was used by a sufficient number of children to make a comparison meaningful. Miss Lee

TABLE 26

AVERAGE PER CENT OF THE DIFFERENT PARTS OF SPEECH USED

	In Chinese	Filipino By rural group	Per cent of interjections used in		Per cent of interjections used in both languages when weighted according to proportion of each language used
Nouns	49.8	59.3	Japanese	6.5	8.3
Pronouns	7.8	4.8	Koreans	1.2	3.1
Verbs	19.2	17.7	Portuguese	0.7	2.1
Modifiers	11.3	10.9	Hawaiians	3.8	3.3
Connectives	1.2	0.6	City Filipino	6.5	4.9
Interjections	6.2	4.7	Rural Filipino		4.2
Enclitics	3.8	Articles 2.0	Chinese		6.7

(16) had analyzed the Hawaiian words in a somewhat different manner, and the few Korean and Portuguese words were translated by *N. Y. A.* assistants.

Most of the words in the three last-named languages were nouns, as is true of the Chinese and Filipino words. It would appear that the less a child uses of a second language, the higher the per cent of nouns. This high per cent of nouns is partly due to proper nouns and objects for which there is no name in English, partly to a lack of opportunity to learn an English word for some family relationship or article about the home. So overwhelming is the proportion of nouns in the non-English language that no other part of speech in it can compete in percentage with the corresponding one in English, except for the interjections. The parts of speech, except for nouns, do follow the same order of frequency in the non-English language as in English where a corresponding part occurs.

The superior maturity of the Caucasian monoglots' speech would be even more marked were the other languages to be considered for the use of nouns instead of pronouns, and scarcity of connectives and articles would be greater. Interjections would be slightly less frequent for some groups, the only criterion of maturity of which that would be true.

Although there is such scarcity of words in some linguistic groups, it seemed worthwhile to see what the proportion of interjections would have been had the child used only English. So the per cent in each language was weighted according to the proportion of each language used. The per cent is still higher than that found for English alone for the most bilingual groups than it is for the Caucasians or for the less bilingual groups.

G. SEX DIFFERENCES IN THE MASTERY OF ENGLISH

Throughout the study, averages were calculated for boys and girls separately, but no significant differences were found, even on those measures that show maturity of speech in general rather than progress in English only.

In the proportion of English, whether of words or of sentences used, the Chinese and rural Filipino girls use a little more and the Japanese a little less than the boys, the averages of the two sexes in the other four groups being within two per cent of each other.

In error in English usage, the girls in six groups are very slightly better than the boys, but the Japanese girls do not do so well as the boys, the difference being almost significant.

In sentence length, the Japanese girls use slightly shorter sentences, the Chinese boys and girls tie, the girls in the other five groups have the advantage. The girls of all groups combined use about 10 per cent more complex sentences.

None of the above differences are statistically significant, although the girls do appear to be slightly more accelerated in speech. The difference, however, does not seem to be as great as that found in previous studies. Since in many cases, especially among the Japanese, the ancestral language is learned more from the home and English outside the home, the fact that even little boys are allowed more freedom from home may be responsible in part for the lower acceleration in the English language found among the girls in Hawaii.

Re-examination of our data showed that although the average age of the sexes in each group was nearly the same, yet in the case of the two-year-olds in every group but one, the boys averaged a little older, the difference being from a third of a month to three months older. Although in each group at some of the latter age levels, girls were a little older, so that the difference in ages between the sexes was at least almost eliminated, yet it is only at a very early age that a month's increase in age will produce an easily measurable improvement in speech.

Therefore, the two-year-olds were carefully paired by age so that in no case was the difference greater than one month. It was possible to match from seven to twelve pairs of boys and girls in

TABLE 27

COMPARISON OF TWO-YEAR-OLD BOYS AND GIRLS AS TO PROGRESS IN SPEECH

	Japanese	Chinese	Filipino Rural	Filipino City	Korean	Hawaiian	Portuguese	Average of the 7 groups
								Total
Number of pairs	12	9	10	7	9	12	10	69
Age in months								
boys	25.5	22.7	24.9	25.3	25.1	26.2	24.9	24.9
girls	25.4	22.7	24.9	25.3	25.1	26.2	25.0	24.9
Average number of sentences								
in English								
boys	18.2	29.4	31.0	34.9	40.9	40.2	48.9	34.8
girls	22.6	36.0	40.5	34.1	41.3	45.0	49.1	38.4
mixed								
boys	7.5	2.2	5.3	6.1	1.4	3.7	0.4	3.8
girls	8.3	1.6	2.1	7.1	1.2	3.2	0.8	3.5
incomprehensible								
boys	2.8	6.7	2.7	3.4	4.6	1.8	0.1	3.2
girls	0.0	1.7	0.4	1.4	1.2	1.2	0.0	1.2
Average error index								
boys	748	799	741	672	595	722	486	680
girls	722	676	825	705	599	586	454	652
Average per cent of English words								
boys	46.8	74.1	75.4	82.4	93.2	89.5	98.5	80.0
girls	49.8	76.2	85.7	82.7	91.5	95.5	99.2	82.9
Average sentence length								
boys	1.48	1.51	1.50	1.69	2.00	1.78	2.43	1.77
girls	1.85	1.56	1.38	2.03	2.18	2.38	2.66	2.00

each group, making a total of 69 pairs. The average per cent of English words, number of English and mixed sentences, error index, and sentence length were calculated for each racial group, and the results are shown in Table 27.

The girls show a distinct superiority on most measures. With the exception of the Koreans, they use more English words and, excepting city Filipinos, more sentences entirely in English—the average difference for all groups being, however, only 2.9 per cent more English words and 3.6 more English sentences.

There is no exception for any groups in the girls' superiority as shown by fewness of incomprehensible remarks—the average number being two less than for the boys. As the number of such remarks recorded is few, this difference means that the boys use almost three times as many incomprehensible sentences as do the girls. In sentence length, too, the other measure, not only of English but of both languages, the girls are in advance of the boys, with the exception of the rural Filipinos. The average difference between the sexes is .23 of a word, or the girls' average is 13 per cent greater than the boys'.

In correct use of English, the girls have but a small advantage; in both groups of Filipinos, the girls make more errors, and the Korean girls are practically tied with the boys. In mixed sentences there is no consistent difference: Japanese, city Filipino, and *Portuguese girls; Chinese, rural Filipino, and Hawaiian boys,* use the most mixed sentences, while the difference between sexes in the case of Koreans is negligible. The percentage of mixed sentences, however, was found to increase and then decrease in number, so it is probably not an appropriate measure of progress at two years.

This analysis of two-year-olds when the sexes are equated exactly as to age suggests that the difference found in other studies as to early precocity of girls as to speech does exist to some extent in the groups studied.

H. HOME INFLUENCE REFLECTED IN THE CHILDREN'S SPEECH

1. *Effect of Parental Birthplace*

It seemed probable that the amount and correctness of English used by the child might be greater in the case of those whose parents were native born than of those whose parents were immigrants.

Almost all the Filipino parents had been born in the Philippines,

TABLE 28

COMPARISON OF THE PROGRESS IN MASTERY OF THE ENGLISH LANGUAGE OF THREE DIFFERENT GROUPS ACCORDING TO BIRTHPLACE OF PARENTS

	Parents born in Hawaii	Language rating of home	Yrs. of English education	No. boys	No. girls	Age	Sentence length	English words	Per cent of sentences that are English	Per cent of sentences that are Mixed	Per cent of sentences that are Japanese	Error index
Japanese	0	1.5	1.0	33	24	50.4	2.9	46.2	37.2	29.5	32.4	554
	1	2.0	4.6	16	15	44.5	2.5	44.2	33.9	23.2	39.4	566
	2	2.5	9.0	17	20	46.1	2.7	60.5	52.0	21.6	25.0	558
Chinese	0	1.6	1.7	12	8	54.8	2.9	72.8	72.8	5.4		458
	1	2.0	4.3	16	7	42.7	3.1	73.2	64.0	4.5		424
	2	3.0	10.0	41	41	46.4	3.5	84.2	83.2	3.0		388
Korean	0	2.1	1.3	45	28	48.0	3.1	94.7	88.4	4.8		463
	1	2.9	7.0	11	15	48.8	3.2	97.8	94.8	2.2		444
	2	2.8	10.1	8	18	46.0	3.1	97.1	92.2	5.4		389

practically all the Hawaiian and Portuguese in Hawaii; but in the case of three groups, there was a considerable difference among the parents as to birthplace.

The children of these three groups were separated into three subgroups according to whether both, one, or none of their parents had been born in Hawaii.

In order to avoid the influence of an uneven distribution of ages in the subgroups, averages were found separately for each of the five age levels, and then these five averages were added together and their average taken. Otherwise, in the smaller groups, the presence of a higher per cent of two-year-olds would have unduly affected the average of that particular subgroup (see Table 28). As was to be expected, the children, both of whose parents were born in America, use more English words and sentences than those, neither of whose parents was so born. Those, one of whose parents had been born in the Islands, hold a different rank in each group as to amount of English used.

The Chinese children also use longer sentences, fewer mixed sentences, with fewer errors, exactly in accord with the number of their parents born in the United States. Differences for the other two racial groups are not so consistent. The only criterion on which the Korean children of native-born parents excel is in amount of error. On the other criteria, those with one parent born in Korea rank first. The apparent superiority of the children whose parents were born in Japan on some criteria is probably accounted for by the advantage of about five months in age that this group has over the other two groups.

2. *Effect of Parental Occupation*

Previous studies have considered the relation of parental occupation or socio-economic status to the child's progress in speech (19, 30, 31). An attempt was made to do the same in this study. As most of the fathers' occupations had been rated on the Barr Scale, the children were divided into three groups: those whose fathers' ratings were 11 and above, small shop-owners, clerical and professional; those whose fathers' occupations were rated from 9:00 to 10:99, mainly skilled trades, sales clerks, etc.; and those whose fathers rated below 9.

As almost all rural Filipinos were plantation laborers, it was not worth while to study this factor in their case; but the average error index, sentence length, and per cent of English words for each of the other six groups were calculated for each of the three occupational groups described above, and the results are shown in Table 29.

Despite the fact that occupational rating and years of English education correlated so low for the Asiatic groups (see Table 4), there is some difference in the amount of English used among the

TABLE 29

AVERAGE PROGRESS IN ENGLISH OF CHILDREN GROUPED ACCORDING TO PARENTAL OCCUPATION

	Japanese *M*	*F*	Age	Parental education	Language rating	Per cent English words	Error index	Sentence length
Barr Ratings								
11:00 up	10	12	47.6	6	2.4	61.5	438	3.17
9:00-10:99	21	27	47.4	4	1.8	46.0	582	2.62
below 9:00	35	20	47.6	2	1.8	50.9	552	2.76
	Chinese *M*	*F*						
11:00 up	30	27	47.7	9	3.0	81.0	404	3.15
9:00-10:99	20	12	48.2	6	2.4	81.8	420	3.03
below 9:00	19	17	48.7	2	1.7	70.0	436	3.10
	City Filipino *M*	*F*						
11:00 up	4	4	48.8	10.9	3.0	92.3	343	3.75
9:00-10:99	0	5	47.8	6.8	2.3	93.2	405	3.75
below 9:00	35	27	47.8	3.9	2.1	90.9	449	3.52
	Korean *M*	*F*						
11:00 up	4	13	47.7	9.6	3.0	95.9	412	3.66
9:00-10:99	11	15	48.4	3.7	2.6	95.6	449	3.20
below 9:00	43	32	47.8	3.5	2.3	95.8	482	3.06
	Hawaiian *M*	*F*						
11:00 up	5	10	48.7	10.8	4.1	97.5	343	3.13
9:00-10:99	4	7	49.1	11.5	3.8	94.6	471	3.19
below 9:00	46	45	48.4	7.4	3.0	95.4	447	3.06
	Portuguese *M*	*F*						
11:00 up	8	4	48.3	8.9	3.3	98.3	353	3.25
9:00-10:99	6	10	49.3	7.4	3.9	98.7	306	3.75
below 9:00	40	38	48.2	5.5	3.0	97.6	397	3.23

children of different occupational levels, except in the case of the Koreans. The difference, however, is very slight in the Hawaiian (many of whom were on relief) and Portuguese groups. In four of the six groups, the amount of error increases as the rating of fathers' occupation decreases. The Japanese children whose fathers have the highest Barr rating rank first both in making fewer errors when speaking English and in length of sentences; but the children whose fathers' rating is lowest on the Barr Scale are second in every one of the three measures. The language rating of their homes is as high as that of the middle group. Perhaps this is due to the fact that with so large a Japanese population as is found in the Islands, it is not necessary for a man to speak English so well in order to hold a good position as is true in the case of some other races.

The Portuguese children also do not show an increase in amount of error with decrease in average Barr rating of fathers' occupation. The first place on all three measures is held by the children of the middle occupational group and second place by those of the highest. But the average age of the children and language rating of the home is highest for the middle occupational group.

In sentence length, the longest sentences are used by the children of either the highest or next highest occupational level, and the shortest by one of the two lower levels, the averages of all the six groups taken together following the occupational ranks; but the differences are slight. Evidently other factors are more important in Hawaii in determining the child's progress in speech than the intelligence of the father as measured by the Barr rating. Probably this measure is not of so much value in the Islands, where the occupations open to men are to some extent determined by their race and the recency of immigration of their racial group (2, 17, 25).

3. *The Effect of Degree of Hawaiian Blood*

The Hawaiian group is quite heterogenous, not quite one-fourth being as much as seven-eighths Hawaiian. Therefore, the Hawaiians were sub-divided into five categories, as shown in Table 30, to see what difference in language development might be found among the different strains. The Caucasian-Hawaiian homes had the advantage in English education and language rating. Nevertheless, in amount of English, scarcity of mixed sentences, and sentence length, the group

TABLE 30
COMPARISON OF CHILDREN OF DIFFERENT MIXTURES OF HAWAIIAN BLOOD AS TO PROGRESS IN SPEECH

	7/8 & 8/8 Hawaiian	Caucasian-Hawaiian	Filipino-Hawaiian	Asiatic-Hawaiian	Of more than two racial stocks
Parents					
Per cent born in U. S.	100	98	58	100	100
Years of English education	7.0	8.6	5.1	7.9	7.9
Language rating	2.5	3.6	2.1	2.6	3.2
Children					
Number boys	19	21	6	5	9
girls	10	28	4	5	18
Age	48.4	48.5	46.4	47.2	49.1
Per cent of English words	94.3	96.2	96.1	95.6	97.2
Per cent of English sentences	88.6	92.8	87.2	88.8	95.0
Per cent of mixed sentences	7.3	4.3	9.0	2.5	2.7
Sentence length	2.9	3.3	3.7	3.2	3.6
Error index	476	402	420	458	418
Sum of ranks on five criteria	23	11	17	16	8

of more than two racial stocks excels them. Only in fewness of errors are the Caucasian-Hawaiians first.

The few Filipino-Hawaiians, despite their disadvantage in home language rating, parental education, and two months' lower age average, use the longest sentences of all and make almost as few errors as does the group of more than two racial stocks. Did they not confuse their three languages so badly, they would have held a higher place.

The nearly full-blooded Hawaiians make the poorest showing of all. All the other groups have come from parents or grandparents of from two to five linguistic stocks. Probably their parents, desiring to be understood by each other and not so desirous of maintaining a non-English language as those people who have mated within their own racial groups are likely to be, have not used much of a non-English language. This would seem to be particularly true of cases where more than two racial stocks have been united in one marriage.

Of course, our cases are very few in number and the differences are not large. It would be interesting to extend this study further.

4. *Effect of Parental Education*

More clearly marked differences are found when the children are grouped according to the English education of the mid-parent. In interpreting Table 31, high school means graduation from high school and above; grade school, completion of eight to eleven years of school; primary, completion of four to seven years; and *L*, one to three years of English schooling. With the higher degrees of education are found quite consistently higher language ratings of home, more parents born in the United States, and higher Barr ratings. Average scores are found as before by averaging the average scores found at each age level. In some racial groups where at one or more age levels no children were found to represent one educational category, those whose parents had completed high school were grouped with those who had completed only eight grades. Data as to education were not complete for all groups. No Hawaiian parents had completed less than four years of English school, and in only six rural Filipino families had the parents completed more than a primary education. Therefore, certain categories are omitted for these groups.

With the exception of the Portuguese and Koreans, the proportion of English words increases for all, and of English sentences, for all but the Chinese, with the increase of parental English education. Mixed sentences show almost an opposite trend. Error index decreases with increasing parental education, except in the case of the Koreans and rural Filipinos. But the Korean children of high school parents average much younger than those in other categories, and the differences in amount of English used for Portuguese and of errors made by rural Filipinos are very small. In length of sentences, a tendency, but not so consistent a one, is found toward the use of longer sentences by children whose parents had more education.

In general, we find considerable correspondence between English education of mid-parent and the child's progress in speech, more especially in those measures used on English speech alone.

5. *Effect of Language Used in the Home*

Although the correlations between parental education, occupation,

TABLE 31

COMPARISON OF THE PROGRESS IN ENGLISH MADE BY THE DIFFERENT GROUPS WHEN DIVIDED ACCORDING TO NUMBER OF YEARS OF ENGLISH EDUCATION OF PARENTS

	Parental				Children					Average per cent of sentences that were			
	Lowest type of school finished	Language rating of home	% born in U. S.	Barr rating	No. of boys	No. of girls	Age in mos.	Length of sentence	English words	English	Mixed	Japanese	Error index
Japanese													
	H	3.7	100	9.82	4	2	41.3	3.41	83.51	79.61	11.0	8.0	514
	G	2.7	100	8.51	7	12	48.5	2.63	63.42	52.42	25.0	22.6	564
	P	2.1	67	8.65	13	11	48.6	2.72	46.43	37.23	34.4	27.6	543
	L	1.7	14	8.24	42	34	47.4	2.63	44.84	36.44	36.0	27.2	574
Chinese													
	H	3.9	98	11.41	11	13	48.2	3.51	88.61	85.42	9.0		517
	G	2.9	96	10.32	17	18	46.9	3.03	88.42	86.71	4.2		400
	P	2.3	83	8.87	14	9	47.2	3.22	80.33	76.73	6.4		408
	L	1.6	39	8.67	26	18	48.7	2.94	71.54	67.84	9.6		440
Filipino													
Rural													
	P	2.0	3	5.05	18	13	46.7	2.92	83.31	71.21	21.2		586
	L	1.9	0	6.03	45	43	47.6	3.01	78.62	63.22	19.5		576
City	G&H	3.0	16	8.51	8	11	48.9	3.71	95.21	89.01	8.9		420
	P	2.0	5	6.50	19	22	45.5	3.52	93.02	80.02	11.8		482
	L	2.0	0	5.89	36	29	47.1	3.53	90.13	79.53	15.5		479

TABLE 31 (*continued*)

	Lowest type of school finished	Parental: Language rating of home	Parental: % born in U. S.	Parental: Barr rating	Children: No. of boys	Children: No. of girls	Children: Age in mos.	Length of sentence	English words	Average per cent of sentences that were: English	Mixed	Japanese	Error index
Korean	H	3.9	83	10.50	5	10	45.0	3.21	95.73	94.42	2.8		434
	G	2.9	74	9.07	11	16	47.6	3.21	98.21	96.01	1.8		427
	P	2.4	38	7.07	9	12	46.9	3.15	96.42	92.63	4.0		425
	L	2.0	4	6.99	36	25	48.4	3.04	94.84	88.04	4.6		471
												% Hawaiian blood	
Hawaiian	H	4.7	100		2	10	49.9	3.41	97.51	94.41	3.3	43	319
	G	2.9	96		40	31	48.7	3.12	97.02	94.22	4.2	65	413
	P	2.9	90		18	21	47.6	2.93	92.73	85.43	8.0	58	509
Portuguese	G&H	3.7	100	8.87	16	14	48.9	3.51	98.33	96.23	3.0		352
	P	2.9	82	6.35	33	34	48.0	3.32	98.71	97.02	2.6		387
	L	2.5	97	4.68	10	6	48.0	3.33	99.81	97.61	2.2		394

H—High school.
G—Grade school.
P—Primary.
L—Less than 4 years.

TABLE 32

COMPARISON OF THE PROGRESS IN ENGLISH MADE BY THE DIFFERENT GROUPS WHEN DIVIDED ACCORDING TO LANGUAGE RATING OF HOMES

	Home language rating	Parental Extent of English education	Parental Barr rating of fathers' occupation	% born in U. S.	No. boys	No. girls	Av. age in mos.	S.L.	% English words	English sentences	Mixed sentences	Error index
Japanese												
	1	89% L	8.66	9	23	14	48.8	2.6	44.8	33.4	26.6	597
	2	58% P	8.17	46	29	35	46.2	2.6	43.4	34.0	25.7	577
	3	41% G	8.90	52	10	7	51.2	2.9	72.8	50.4	18.1	467
	4&5	100% H	11.74	79	4	3	46.7	3.9	87.5	85.8	11.6	266
Chinese												
	1	81% L	8.85	52	13	8	50.4	3.15	45.8	45.4	12.2	428
	2	56% P	10.31	61	26	17	45.4	3.03	79.7	76.5	9.0	455
	3	77% G	10.26	87	18	17	45.8	3.05	91.5	86.8	6.3	377
	4&5	65% H	11.30	98	13	13	48.6	3.48	92.6	88.0	7.9	305
Filipino Rural		years										
	1	1.4	5.27	00	9	5	46.6	2.98	47.3	23.2	23.1	641
	2	3.4	5.58	02	56	54	47.4	3.00	83.3	70.2	17.6	582
City												
	2	4.0	5.89	07	58	55	47.6	3.34	90.9	81.2	14.4	493
	3&4	10.3	8.96	12	5	7	48.9	3.91	94.4	91.3	10.0	392
Korean												
	2	1.8	6.84	13	39	36	52.0	3.01	95.2	89.4	4.9	468
	3	9.2	7.86	71	15	21	47.4	2.74	96.2	91.6	3.5	420
	5	11.4	10.47	43	3	4	47.2	2.99	99.4	94.4	0.8	306

TABLE 32 (*continued*)

Home language rating	Parental: Extent of English education	Parental: Barr rating of fathers' occupation	% born in U. S.	No. boys	No. girls	Av. age in mos.	S.L.	% English words	English sentences	Mixed sentences	Error index
Portuguese											
2	4.6	5.32	83	21	10	48.6	3.30	98.4	96.4	3.2	423
3	6.3	6.75	98	21	28	48.1	3.30	98.6	97.4	2.4	395
4&5	8.4	8.86	100	15	10	49.2	3.46	98.6	97.0	2.4	311
Hawaiian											
		% Haw. blood									
2	7.0	70	92	27	21	48.7	3.08	95.8	91.0	6.1	450
3	8.1	59	95	16	15	48.3	3.20	97.8	95.3	6.4	411
4	8.0	72	100	11	13	48.5	3.16	97.2	92.4	4.5	459†
5	10.0	42	100	8	14	49.0	3.24	97.2	95.6	2.8	319

L—Little or no English education.
P—Primary but not grade school.
G—Grade school, but not completed high school.
H—High school graduate or more.
†Excessively high score at 2 years.

and the language rating of homes are not high, yet when the different groups were separated into sub-groups on the basis of the homes' language rating, the amount of English education and of per cent born in the United States of parents, and, in most cases, the Barr rating of the fathers' occupation, increased with higher language rating for each group. Sub-groups for which any age-level did not have at least one representative were either omitted or were combined with another group. The averages of each age-level were found, and then the mean of the five averages was calculated as before in order to discount the effect of an uneven scattering throughout the age levels of children in different sub-groups. The results of this classification are shown in Table 32.

The most marked differences are in the more bilingual groups in amount of English used, whether measured by words or sentences. The less bilingual groups show very little difference in this respect. In every racial group, however, there is, with two exceptions, a steady decline in the number of errors with increase in the home's language rating. The exceptions are from rating one to two for the Chinese children (probably due to five months' advantage in age for the former); and a higher score for the group rated four than for those rated lower among the Hawaiian children, probably due to the excessively high average found for the few two-year-olds at that level.

The number of mixed sentences decreases with higher language rating. There is no decrease after level three for Portuguese and a slight gain after that point for the Chinese; and the percentages at levels two and three are almost identical for the Hawaiian. In sentence length, the only clear difference is that, at levels four and five taken together, the longest sentences are found; and, when a comparison is possible, sentences at level five average a bit higher than at level four. Levels one and two show either no difference or, in the case of the Chinese, slight loss from level one to two, which might be on account of the age advantage at level one again.

The average language rating of homes is reflected in the use of English names for the children. Many families here, as is evident in the birth columns of the vital statistics in the newspapers, give their babies an English name as well as another name in the language of their racial ancestry; but almost no families supplied the observers with more than one name. English or Anglicized names only were given for all the Portuguese children, 94 per cent of the Hawaiian, 79 per cent of the Korean, 68 per cent of the Chinese, 36 and 32 per cent respectively for city and rural Filipino, and 30 per cent of the Japanese. Except for the Filipinos, the per cent of English names in each group correlates perfectly in rank with the proportion of English spoken and is in quite close agreement with the rank of the homes according to language rating.

6. *Effect When Parents Speak Different Dialects*

Two children for whom records were taken but not used both had parents of different racial ancestry. These children used more English than the average children of their age in either parental stock. As it seemed probable that in such homes the parents would be more likely to need to use English as a medium of conversation, the Filipino records for which data as to parental dialect were available were analyzed to discover if there were any difference between children whose parents spoke the same and those whose parents spoke different dialects. The results are shown in Table 33. The 21 pairs of children were matched exactly as to sex and age. The average advantage for the group whose parents spoke different dialects was, in English education, about three months and 0.27 of a point on the Barr Scale. As both groups were rated two or more as to language rating, both sets of parents used some English in the home, although it was highly "pidgin." The difference in rating was in favor of the different dialect group, but that was only two-tenths of a point. However, only 93 per cent of this group were born in Hawaii, 98 per cent of the other; so it otherwise would not have been expected that the parents using different dialects would have had even a slight advantage in the language rating. Their children do not use as long sentences, but do use a little more and slightly more correct English than do the Filipinos whose parents use the same dialect.

7. *Effect of Order of Birth*

In order to study the effect of birth order on mastery of English, two methods were used. By the first method (Table 34), as many as possible of the children of each racial group were paired, a first- or second-born child with a later-born child of the same sex and of approximately the same age and home background.

TABLE 33

A COMPARISON OF THE PROGRESS IN MASTERY OF THE ENGLISH LANGUAGE BY FILIPINO CHILDREN WHOSE PARENTS SPEAK DIFFERENT FILIPINO DIALECTS WITH THOSE WHOSE PARENTS SPEAK THE SAME DIALECT

	Parental								Per cent of sentences that were		
	English education	Home language rating	Barr rating	% born in Hawaii	No.	Age in mos.	Sentence length	English words	English	Mixed	Error index
Parents speak Different dialects	5.3	2.2	7.07	93	21	38.7	2.68	91.3	83.7	10.2	559
The same dialect	5.0	2.0	6.80	98	21	38.7	2.90	90.5	80.5	12.4	597

TABLE 34
COMPARISON OF PROGRESS IN THE ENGLISH LANGUAGE MADE BY CHILDREN WHO WERE EITHER THE ELDEST OR NEXT ELDEST IN THEIR FAMILIES WITH CHILDREN WHO WERE THIRD OR LATER BORN

					Filipino									
	Chinese		Japanese		rural		city		Korean		Hawaiian		Portuguese	
	O	Y	O	Y	O	Y	O	Y	O	Y	O	Y	O	Y
Number	26	26	27	27	40	40	39	39	25	25	37	37	37	37
Boys	15	15	10	10	21	21	20	20	12	12	20	20	19	19
Girls	11	11	17	17	19	19	19	19	13	13	17	17	18	18
Parental education	7.0	6.8	4.6	4.5	3.2	2.8	3.7	3.5	5.3	4.9	7.5	7.5	6.5	5.9
Language rating	2.4	2.7	2.0	2.0	1.9	1.9	2.1	2.1	2.4	2.4	3.0	3.0	—	—
Barr rating	9.63	9.85	8.86	8.64	5.27	5.32	5.97	6.22	6.65	7.38	—	—	7.05	6.81
Per cent born in Hawaii	79	79	54	54	3	0	3	0	52	20	100	100	100	100
Per cent of English words	73.2	85.1	37.1	49.8	78.0	79.4	91.9	92.6	96.6	94.7	95.4	95.9	96.6	94.7
English sentences	69.4	83.8	32.7	40.5	59.8	66.4	79.4	80.7	93.9	88.6	91.3	92.3	93.9	88.6
Mixed sentences	8.6	5.4	31.4	23.3	22.0	17.5	15.4	14.3	4.6	3.7	4.9	5.4	4.6	3.7
Sentence length	3.1	3.2	2.80	2.74	2.90	3.06	3.5	3.4	3.3	3.2	3.2	3.0	3.3	3.2
Error index	408	382	558	569	608	597	472	474	470	428	433	432	470	428
Average age of children	49.5	49.5	51.8	51.9	46.4	46.3	48.3	48.3	48.4	48.5	49.1	49.1	53.4	53.5

O—Two oldest children in families.
Y—Younger members of families.

The families of our samples were rather large, so the birth-order of the later-born children ranged from three to thirteen. The second-born children had either not had the advantage of an elder sibling in school or, in a very few cases, for only a few months' time.

In matching children, the average age of each paired group differed by not more than one-tenth of a month, parental education by not more than six-tenths of a year, in Barr rating in only one case by more than three-tenths of a point. Only in the case of the Chinese was there any difference in language ratings, and only in the case of the Koreans was there any difference greater than three per cent as to parental birthplace. Data were lacking in too many cases of otherwise usable pairs as to language rating for the Portuguese and satisfactory Barr rating for the Hawaiians (too many of whom were on *F. E. R. A.* projects).

In the case of those groups who, as a whole, use a considerable proportion of language other than English, the younger children have the advantage as to amount of English used and fewness of mixed sentences, and two of these groups use slightly longer sentences. On the contrary, among the younger children for the four groups using very little but English, there is very little difference on these measures, but the larger of the differences that do occur are in favor of the elder children. As to errors, in four cases the elder children make an average of 30 more per thousand words; while in the other three cases, the younger average five more errors per thousand words.

The second method was to compare the progress of older and younger members of the same family by finding the percent of the racial average for his age attained by each child (see Table 35). According to this method, the younger children use more English words in the three most bilingual groups, the next two use about the same, and in the case of the Portuguese and Hawaiian they use less. The same is true as to English sentences except for the Chinese. As mixed sentences do not follow a straight-line curve, this comparison as to their use for children whose average age differs is not practicable. Except in the case of the Filipinos, the older children use longer sentences. The elder children for their age make fewer errors in the case of the Portuguese and Hawaiian and very slightly less in the case of the Japanese and Korean, the reverse being true for the other groups.

TABLE 35

OLDER AND YOUNGER BROTHERS AND SISTERS COMPARED AS TO PER CENT ATTAINED OF AGE-NORM FOR OWN RACIAL GROUP ON FIVE CRITERIA OF LANGUAGE DEVELOPMENT

Number of sibling pairs of	Japanese	Chinese	Filipino Rural	Filipino Honolulu	Korean	Portuguese	Hawaiian
	16	28	28	25	38	24	7
Older							
Average age in months	64.4	60.8	61.1	60.0	61.5	60.2	71.6
Average per cent of norm reached in error index	106.1	106.4	103.5	105.7	96.0	99.8	84.0
Sentence length	112.8	101.5	95.6	103.0	100.4	103.1	115.6
Mixed sentences	98.2	99.2	106.4	118.2	39.1	46.8	113.0
Mixed sentences*	112.3	112.3	—	—	63.7	91.3	—
English sentences	100.5	101.3	90.4	97.0	102.0	99.2	97.6
English words	99.0	93.1	99.0	99.2	100.8	99.9	101.0
Younger							
Average age in months	38.1	36.6	40.9	36.9	37.1	39.0	50.9
Average per cent of norm reached in error index	111.8	103.8	103.5	103.7	99.3	111.2	109.1
Sentence length	111.1	95.3	103.3	112.4	95.9	97.6	98.6
Mixed sentences*	91.8	91.0	137.4	137.8	37.3	54.3	94.7
	93.4	110.8	—	—	67.6	100.3	—
English sentences	116.1	100.6	96.4	98.8	101.4	98.5	101.3
English words	111.1	100.3	101.5	101.6	100.0	96.6	100.3

*The figures in this column are the averages found when all children who use no mixed sentences are excluded.

In general, it would seem that the elder children have a similar advantage in mastering language to that which has been found in studies of monoglots, due to greater adult attention, in those groups where little but English is used; but where another language is much used, the younger children in their preschool years learn to use more English from their elder siblings and are able to separate the two languages a little better. On other measures, however, they appear to have no advantage in speech.

I. OTHER ENVIRONMENTAL FACTORS

1. *Effect of Place of Residence*

a. In city or country. In the preceding chapters, two Filipino groups have been constantly compared. By every criterion used, the city children have the advantage in mastery of English. They use more English at each age, make fewer errors, are ahead in the elimination of mixed sentences, and use longer sentences. In the foregoing chapter, it was shown that this is true when groups equal in amount of parental education are compared.

Our rural cases were found on sugar plantations. In Hawaii, the plantation laborers live in camps or small villages rather than in scattered farmhouses. Oftentimes those assigned to a particular camp are all of the same racial ancestry. Sixty-six per cent of our rural subjects lived in Filipino camps. The children in such camps did not have to learn English or any other than a Filipino dialect in order to be understood by their playmates.

Mr. Masuoka, in the course of his study of the Americanization of Japanese in the Islands, recorded the conversations of 17 rural Japanese children according to our method. These were matched according to age, sex, parental education, and Barr rating of fathers' education with 17 of our Japanese subjects from Honolulu. These rural children use 21 per cent less of English words, average almost 14 fewer English sentences and 10 more mixed sentences, and make 78 more errors per thousand English words than do the city children (see Table 36); but, although their complex and compound English sentences are fewer, their sentences average longer when all 50 sentences are considered. They ask only a little more than half as many questions, and the egocentric indices of the two groups are almost the same. Many more of their English words are conjugated,

TABLE 36

COMPARISON OF RURAL JAPANESE CHILDREN WITH JAPANESE CHILDREN LIVING IN HONOLULU

	Rural Japanese children	Japanese children in Honolulu	Difference
Number of children	17	17	
Barr rating of fathers' occupation	8.33	8.38	— .05
Years of parents' schooling	*6.6*	*6.4*	*0.2*
Age in months	51.2	51.2	0.0
Average number of words spoken			
in English	67.4	88.3	—20.9
in home language	133.0	60.4	72.6
in Hawaiian	2.8	1.3	1.5
in another language	0.2	0.0	0.2
Average number of sentences entirely in English	10.1	23.9	—13.8
entirely in home language	15.7	11.7	4.0
in another language	0.4	0.4	0.0
a mixture of languages	23.8	14.1	9.7
Error index	600.9	522.6	78.3
Sentence length	3.5	2.9	0.6
Number of complex and compound sentences	3.0	3.4	— .4
Number of questions	3.9	6.9	— 3.0
Egocentric index	334.2	348.2	—14.0
Per cent of verbs conjugated	207.7	81.1	126.6
Per cent of English words that were			
copulas	0.3	0.7	— 0.4
connectives	1.1	2.6	— 1.5
articles	0.8	2.0	— 1.2
nouns	17.5	23.2	— 5.7
pronouns	23.2	20.8	2.4
interjections	6.6	9.7	— 3.1
verbs	18.3	19.3	— 1.0
modifiers	31.8	22.4	9.4
Per cent of English words	38.1	59.1	—21.0

mainly due to their overuse of the past form of the single verb *"broke."* They use fewer copulas, connectives, articles, and nouns, and less frequently use an English word when resorting to interjections. Altogether they, like the rural Filipinos, show much less mastery of English than do the city children.

Miss Lee (16) studied a few rural Hawaiian children in the course of her investigation of the use of Hawaiian words in the Islands. Although she observed all the children in one rural community on Oahu, where most of the inhabitants were Hawaiian, only four of them were of the age of our subjects. They use more

Hawaiian words and make more errors in the use of English than do comparable cases among our subjects.

b. In areas of mixed population or in those where one racial group was concentrated. As the failure to meet children who cannot speak the non-English home language seems to be the most probable potent factor in the rural children's backwardness in English when compared with city children, a comparison was made of Honolulu children residing in sections of the city populated almost entirely by persons of the same racial ancestry and therefore situated somewhat as the rural children are, with children residing in areas where the population is of diverse racial antecedents (see Table 37). Residential areas of the first type will be referred to hereafter as concentrated and of the second type as mixed. The classification of these areas in Honolulu had been made by the Sociology Department at the University, and Dr. Andrew Lind, head of the department, kindly permitted us access to these data. The Portuguese and Koreans, especially the latter because of the small proportion it composes of the total population, are not found in concentrated areas of any size. For the other four groups, children from the two types of residential areas were matched with each other on the basis of age, sex, parental education, and, except in the case of the Hawaiians for whom the occupational rating was not obtained in too many cases, by the Barr rating of the father's occupation also. The language rating of the home was also considered, no child being matched with another child if the rating differed by more than one point on the scale; but the homes in mixed areas averaged so much higher than those in concentrated areas in this respect that the difference between the pairs was the largest in this criterion of any. To the Honolulu pairs were added 17 rural Filipinos matched by the same criteria, one member of each pair of which lived in a Filipino camp, the other in a mixed camp.

There were only 105 pairs which could be found among the 875 cases. Most of the Hawaiians and city Filipinos lived in mixed areas, most of the Japanese and rural Filipinos in concentrated areas. The Chinese showed the most diversity as to residence in this respect, and it was possible to match half as many pairs of them as for all other four groups taken together.

The average of the five groups from mixed areas was 9.3 per cent more English words and 7.9 per cent more entirely English sentences

TABLE 37

COMPARISON OF PROGRESS IN THE ENGLISH LANGUAGE MADE BY CHILDREN RESIDING IN COMMUNITIES COMPOSED MAINLY OF PEOPLE OF THE SAME ANCESTRY WITH THOSE CHILDREN RESIDING IN COMMUNITIES OF MIXED RACIAL ANTECEDENTS

							Filipino				Av. of 5 groups	
			Chinese		Japanese		City		Rural		Total	
	*C**	*M*†	*C*	*M*	*C*	*M*	*C*	*M*	*C*	*M*	*C*	M
No.	35	35	13	13	18	18	22	22	17	17	105	105
Boys	21	21	10	10	13	13	12	12	9	9	65	65
Girls	14	14	3	3	5	5	10	10	8	8	40	40
Parental education	6.6	7.0	6.6	6.3	6.9	6.9	5.2	5.3	3.3	3.2	5.7	5.7
Language rating	2.1	2.4	2.2	2.7	2.5	2.5	2.1	2.1	1.8	1.9	2.1	2.3
Barr rating	9.54	9.63	9.04	9.19	—	—	6.46	7.50	5.59	5.58	7.66	7.98
Age	47.9	47.6	51.6	51.5	50.1	49.9	55.9	55.9	42.9	42.9	49.7	49.6
Per cent of English words	70.8	85.1	54.2	69.6	93.5	95.8	92.1	94.6	74.9	84.5	76.7	86.0
English sentences	70.9	80.9	39.8	60.8	88.0	94.3	82.5	84.7	59.2	62.4	68.1	76.7
Mixed sentences	9.6	7.7	35.9	21.4	7.8	3.8	13.6	13.4	19.8	22.2	10.9	13.7
Sentence length	3.0	3.1	3.1	3.2	2.7	3.0	3.6	3.6	3.1	2.8	3.1	3.1
Error index	444	408	553	446	480	463	448	455	632	643	511	483

*The scores of children residing in areas where their own racial group is concentrared.

†The scores of children residing in areas of mixed population.

than for those from concentrated areas. The difference was in the same direction for all groups on both criteria. Two groups show more mixed sentences used by those residing in mixed areas; three show less. This difference depends on the amount of English used, the groups using fewer mixed sentences when residing in concentrated areas being those who use least English. Sentence length does not appear to be affected by this factor. Error in use of English is less, excepting for Filipinos, for those residing in mixed areas.

It would appear to be a decided advantage, if a child is to learn a second language besides his mother tongue, for him to reside in a neighborhood where few of the children speak the language spoken in his own home. As that is true of our Korean group as a whole, it is not surprising that they surpass the other bilingual groups in their use of English.

2. *Effect of Kindergarten or Nursery School Attendance on Progress in the Mastery of English*

In order to study the effect of kindergarten or nursery school attendance on mastery of English, two methods were followed: one hundred records were taken at kindergarten, and the records taken at home of children who attended kindergarten or school were compared with those of children who did not.

The records taken at kindergarten were all taken in Honolulu kindergartens and were taken during free play or other periods when the child's conversation was not directed by his teachers. One of the kindergartens used includes a nursery school department, so the age range of these subjects is similar to that of our main groups.

As the home records numbered about 900, it was possible to find among them conversations to match each of the kindergarten records.

The children were paired exactly on the basis of sex and race, were within three months of the same age, their parents had had approximately the same amount of English education, and their fathers averaged within a few points on the Barr Scale for rating occupations on which the average for the kindergarteners' fathers was 10.04 against 9.76 for the other children. Only 62 per cent of the former's parents were born in the United States and 70 per cent of the latter's; but the former's parents averaged 7.9 years of English education against 7.3 for the latter's. As the age difference is but 0.1 month, the two groups are quite well matched (see Table 38).

TABLE 38

COMPOSITION OF GROUP WHOSE RECORDS WERE TAKEN AT KINDERGARTEN AND THAT OF CHILDREN PAIRED WITH THEM WHOSE RECORDS WERE TAKEN AT HOME

	Children				Parents		
Records taken at	Number Boys	Number Girls	Age in months	$\sigma_{dist.}$	Born in U. S. A.	Years of English schooling	Barr rating of fathers' occupations
Kindergarten	49	51	52.4	14.8	62%	7.9	10.04
Home	49	51	52.3	15.2	70%	7.6	9.76

When the conversations of the two groups were compared, it was found, as summarized in Table 39, that the children at kindergarten used almost 30 per cent more sentences in which English was exclusively used. They mixed words of more than one language less than one-fifth as much and made about 15 per cent fewer errors when speaking English. They used a slightly higher number of complex and compound sentences, and their sentences were about 10 per cent longer. As other parts of the study have shown, our bilingual children conjugate very few verbs, use very few copulas, articles, and connectives, and use many more interjections than do monoglots of the same age who speak more perfectly. On every one of these criteria, the children at kindergarten show a much better use of the English language.

Our kindergarten records were taken at Castle and Moiliili. The former kindergarten charges tuition and enrolls a large number of children from homes where English is the only language spoken, but the latter does not. The advantage that a bilingual child has in learning to master English in a school where there is a nucleus of monoglots can be seen by comparing the averages of Japanese children of about the same age in the two kindergartens. In all measures except length of sentence, the children at Castle excel the others, although they average two months younger: but as the children at Moiliili averaged higher in the use of English than did their corresponding mates observed at home, it is evident that attendance at kindergartens where English is spoken, even if none of the children speaks it exclusively, is an advantage to a child's mastery of the language.

To study the amount of carry-over of the use of English to the

TABLE 39

A COMPARISON OF 100 RECORDS OF CHILDREN'S CONVERSATIONS RECORDED AT KINDERGARTEN WITH THE SAME NUMBER RECORDED AT HOME

	Kindergarten attendants	Non-Kindergarten	Japanese pairs at Castle	Japanese pairs at Moiliili	Japanese pairs at Kindergartens
Average age in months	52	52	60	60	
Number of Japanese children	61	61	16	16	
Number of Chinese children	33	33			
Number of Korean children	1	1			
Number of Filipino children	1	1			
Number of Part-Hawaiian children	4	4			
Per cent of English words	97.4	70.8	97.8	95.3	
Number of English sentences	47.2	31.6	48.1	44.1	
Number of mixed sentences	1.7	9.4	1.4	4.4	
Errors per 1,000 words	387	451	329	466	
Number of words per sentence	3.4	3.1	3.7	3.8	
Per cent of verbs conjugated	18.6	15.9	20.6	8.5	
Per cent of English words used that were:					
copulas	3.0	0.7			
articles	2.9	1.7			
connectives	3.9	2.6			
interjections	4.2	6.6			
Number of complex and compound sentences	3.0	2.9			
Number of questions	4.5	6.2			
σ of dist. for					
error index	142	167			
sentence length	0.87	0.87			
σ_p for English words	1.7	4.5			
Number of					
Chinese sentences	0.1	1.0			
Hawaiian sentences	0.8	0.2			
Japanese sentences	0.3	7.6			
Filipino sentences	0.0	0.2			
Incomprehensible sentences	0.5	0.2			

Differences between kindergarten and other children

	Per cent English words	Sentence length	Error index
Difference	26.6	0.3	64
σ of difference	4.8	0.135	23.4
Critical ratio	5.5	2.2	2.7
Chances in a 1,000 of a true difference	1,000	986	997

TABLE 40

COMPOSITION OF GROUPS WHO ATTEND KINDERGARTEN AND OF CHILDREN PAIRED WITH THOSE WHO DO NOT

	Number of		Age in months	Parents		Barr rating	Language rating of home
	Boys	Girls		Born in U. S.	Years of English schooling		
Kindergarten							
Chinese	8	8	58.6	75%	6.3	9.01	2.6
Japanese	4	1	60.6	30%	2.0	8.35	1.6
Korean	7	8	62.9	20%	3.9	7.32	2.7
Filipino	7	7	59.0	0	4.5	6.08	2.1
Total	26	24	60.3	33%	4.6	7.41	2.4
Hawaiians	4	4	68.8	100%	8.3		2.4
Non-Kindergarten							
Chinese	8	8	58.6	81%	7.6	8.96	2.6
Japanese	4	1	60.6	30%	2.2	8.19	1.6
Korean	7	8	62.9	17%	3.9	7.79	2.3
Filipino	7	7	59.6	0	3.5	6.55	2.1
Total	26	24	60.3	34%	4.7	7.54	2.3
Hawaiians	4	4	68.8	100%	8.4		2.4

home situation, the home records were scrutinized, and all those of children who were reported to be in attendance at kindergarten or had started to school were selected. It was possible to find for most of these children mates paired on the same basis as described above (see Tables 40 and 41). Fifty pairs were of children of the bilingual groups and eight of part-Hawaiian ancestry, a group which has been shown above to use, even at two years, over 90 per cent English words.

Considering the 50 bilingual pairs first, it is found that the differences between the groups, although much less, favor the group that attends kindergarten in most cases, whether the racial groups are considered separately or together. Per cent of mixed sentences is the only criterion in which the non-kindergarteners in the combined group excel the others. This exception is found to be due entirely to the Japanese group, who are much less fluent in English than are the others and whose improvement in its use involves a transition from the use of Japanese to mixed sentences before attaining the use of sentences entirely English. The kindergarten children of this group used an average of less than six Japanese sentences when playing with other children about their own homes, while the non-kindergarten Japanese in the same situation averaged about 15 such sentences.

TABLE 41

A COMPARISON OF THE ENGLISH USAGE OF PRESCHOOL CHILDREN OF DIFFERENT RACES WHO ATTEND KINDERGARTEN WITH THOSE WHO DO NOT

	Chinese		Japanese		Korean		Filipino		All Non-Hawaiian		Hawaiian	
	*K**	*N*	*K*	*N*	*K*	*N*	*K*	*N*	*K*	*N*	*K*	*N*
Number of children	16	16	5	5	15	15	14	14	50	50	8	8
Boys	8	8	4	4	7	7	7	7	26	26	4	4
Girls	8	8	1	1	8	8	7	7	24	24	4	4
Average age in mos.	59	59	61	61	63	63	60	60	60	60	69	69
Per cent of English words	91.6	85.1	59.4	55.2	97.9	96.5	95.1	94.0	91.2	88.2	95.4	96.3
Number of English sentences	43.9	42.2	21.8	23.0	47.2	46.6	44.4	42.5	42.8	41.7	46.8	46.6
Number of mixed sentences	4.1	3.8	19.6	11.6	2.2	3.0	5.2	6.8	5.4†	5.3	2.8	2.5
Errors per 1,000 words	314	332	493	486	337	395	410	402	336	386	348	386
Per cent of words per sentence	3.8	3.6	3.5	2.8	3.7	3.6	4.0	4.0	3.8	3.7	3.6	3.9
English words that were:												
copulas	1.2	1.0	0.2	0.2	1.4	1.0	1.5	0.8	1.2	0.8		
articles	3.1	2.2	2.2	0.9	2.9	3.1	2.2	2.3	2.7	2.7		
connectives	4.9	4.0	1.0	1.2	4.5	4.2	3.7	2.3	4.0	3.3		
Number of complex and compound sentences	4.2	3.5	4.8	2.8	3.6	3.1	6.9	5.3	4.8	3.7		

**K*—Kindergarten attendants.
N—Non-kindergarten attendants.
†Excluding Japanese the averages are $K = 7.6$, $N = 9.6$.

The few Hawaiian pairs show very little difference in amount of English spoken, but there is 10 per cent less error on the part of the kindergarten attendants.

In some cases of children not attending kindergarten, they are probably hearing as good English at home as they would at kindergarten. This tends to reduce the difference between the two groups. So, also, does the fact that the time at kindergarten had been very short, averaging less than four months in the case of these 58 pairs, whereas the children whose records were taken at kindergarten had been in attendance in every case, except for a few two-year-olds, for at least six months.

It appears, then, that even a few months' attendance at kindergarten results in so much greater practice in speaking a more nearly correct English on the part of bilingual preschool children in Honolulu, that English is used more and spoken more correctly by children of such attendance not only while they are at kindergarten but also when they are at home.

J. THE EFFECT OF THE PRESENCE OF ADULTS ONLY AS CONTRASTED TO THE SITUATION IN WHICH OTHER CHILDREN ALSO ARE PRESENT ON THE CHILD'S SPEECH

When it was discovered that, through a misunderstanding of directions, one of the Portuguese recorders had taken a number of records of children over two years of age when the child observed had been playing alone with only adults present, it was decided to add a few more records taken under the same conditions to see if a comparison of these records with those taken in the desired situation would show the same contrast as had been found before when comparing monolingual children in the two situations (31).

Twenty-five records were secured under Situation *A*, or that in which the child was alone with adults. Children observed in this situation were paired by race, age, sex, parental occupation, and education, with children observed in Situation *C*, that in which the child studied was playing with other children. The results of this comparison are shown in Table 42. In general, the children use more English, longer sentences, make fewer errors, and use fewer mixed sentences, conjugate more verbs, use more complex and compound sentences, connectives, articles, and fewer interjections in Situation *A*. All these differences are in the same direction as they were in the previous study. Although different methods had been used in the two studies to determine indices of egocentricity, both

TABLE 42

Comparison of Children Whose Conversations were Recorded in Situation *A* with Those Whose Conversations were Recorded in Situation *C*

	All cases Situations			Portuguese and Hawaiian situations			Filipino, Japanese and Korean situations		
	A	*C*	Difference	*A*	*C*	Difference	*A*	*C*	Difference
Number of children	25	25		18	18		7	7	
Average									
Barr rating of fathers' occupation	9.08	8.19	0.89	9.38	8.35	1.03	8.50	7.78	.72
Years of parents' English schooling	8.5	7.9	0.6	9.4	8.9	0.5	7.9	6.6	1.3
Age in months	47.8	47.8	0.0	46.3	46.2	0.1	51.9	51.7	0.2
Average									
Number of words spoken:									
in English	302	156	146	228	170	58	152	120	32
in home language	9.0	3.0	6.0				26.6	39.6	— 13.0
Hawaiian				1.2	1.6	— 0.4	2.0	0.4	1.6
Portuguese				1.2	1.8	— 0.6			
in another language	0.9	0.7	0.2	0.1	0.3	— 0.2	0.6	0	0.6
Average									
Number of sentences:									
entirely in English	44.9	42.4	2.5	48.4	47.1	1.3	36.0	30.1	5.9
entirely in home language	1.3	3.8	— 2.5	0.3	0.8	— 0.5	4.0	11.6	— 7.6
in another language	0.1	0.1	0.0	0.1	0.1	0.0	0.1	0.1	0.0
a mixture of languages	3.7	3.8	— 0.1	1.3	2.1	— 0.8	9.9	8.1	1.8

TABLE 42 (*continued*)

	All cases Situations			Portuguese and Hawaiian situations			Filipino, Japanese and Koreans situations		
	A	*C*	Difference	*A*	*C*	Difference	*A*	*C*	Difference
Error index	239.5	432.2	—192.7	145.1	362.2	—217.1	482.4	612.0	—129.6
Sentence length	4.3	3.5	0.8	4.6	3.5	1.1	3.6	3.4	0.2
Number of complex and compound sentences	3.6	2.9	0.7	3.7	2.7	1.0	3.6	3.4	0.2
Number of questions	5.2	5.5	— 0.3	4.7	5.7	1.0	7.1	4.9	2.2
Egocentric index	539.1	527.1	12.0	560.4	594.5	— 34.1	484.3	353.4	130.9
Per cent of verbs conjugated	311.8	191.5	120.3	355.7	218.9	136.8	199.0	120.9	78.1
Per cent of English words that were:									
copulas	1.6	1.0	.6	2.0	1.3	0.7	0.7	0.2	0.5
connectives	7.6	3.3	4.3	9.3	3.8	5.5	3.3	2.0	1.3
articles	4.9	2.7	2.2	5.9	3.4	2.5	2.2	1.1	1.1
nouns	22.3	18.9	3.4	21.4	19.2	2.2	24.6	17.9	6.7
pronouns	20.8	20.8	0	20.2	22.1	— 1.9	22.3	17.7	4.6
interjections	1.9	3.1	— 1.2	1.7	1.5	0.2	2.3	7.0	— 4.7
verbs	25.5	28.2	— 2.7	25.6	28.8	— 3.2	25.4	26.6	— 1.2
modifiers	17.9	23.2	— 5.3	17.1	21.0	— 3.9	19.9	28.9	— 9.0
Per cent of English words	94.2	88.9	5.3	98.9	97.5	1.4	81.9	69.0	12.9

agreed in finding a slightly higher index in Situation *A*. Only in one comparison do the studies differ: slightly more questions were asked in Situation *C* by the children in Hawaii, whereas many more were asked in Situation A by the group previously studied.

That more English was used in Situation *A* seems surprising; but since 18 of the children studied in this situation were Portuguese and Hawaiian, whose parents speak English almost entirely, its greater use when alone with adults by these children is only in line with the tendency to a more perfect speech when in that situation.

Separating the children observed into two groups, 18 Portuguese and Hawaiian, in whose homes little but English is used, and the 7 Filipino, Japanese, and Korean in whose homes another language is often if not usually heard, it is found that although the difference in the two situations remains in the same direction for both groups on almost all criteria, in respect to number of questions asked it does not.

The former group is found to be the one responsible for the exception in results to the previous study, for the latter group do ask more questions in Situation *A*. As the Hawaiian children in Situation *C* did not ignore the adult observer but talked to her, it is probably due to this fact that the present study does not confirm the previous results.

As to the choice of language, the few cases of Japanese, Korean, and Filipino also used more English when alone with adults. Examination of their records reveals comparatively few remarks addressed to their parents, for in five cases, visitors or relatives who used mostly English were present as well as the observer, and these people were addressed often when the child was not jabbering to himself.

K. RESULTS FROM REPEATED STUDY OF FORTY-FOUR CASES

Forty-four of our subjects were observed twice, the second record being taken about nine months after the first. A comparison of their two records shows the same trend with age found for the groups as a whole (see Table 43). There is an increase with age in the length of sentence, the amount of which is less at each age level. The amount of error is not reduced to any marked extent except from the first and last age levels, while proportion of English, whether measured by words or sentences, increases with age up to six years. At this age level, there are but four cases, so that even one exception to the rule may swing the average. There was one child whose first record, taken at six years, was the best of all our thousand cases.

TABLE 43

AVERAGE YEARLY GAIN IN USE OF THE ENGLISH LANGUAGE MADE BY 44 CHILDREN FOR WHOM TWO OR THREE RECORDS WERE TAKEN

First record	Number of cases	Average interval between records	Gain in sentence length	Reduction errors per 1,000 words	Increase in per cent of			Attending school or kindergarten
					English sentences	Mixed words	English words	
At 2 yrs.	12	10 mos.	1.73 words	—232	17.4	— 2.4	11.2	4
At 3 yrs.	15	10 mos.	.72 words	—098	9.6	—11.5	2.7	3
At 4 yrs.	5	11 mos.	.31 words	—014	0.6	0.0	1.4	2*
At 5 yrs.	8	8 mos.	.18 words	—027	5.7	— 9.9	8.2	4*
At 6 yrs.	4	6 mos.	.10 words	—212	1.2	4.0	16.0	4*

*Both children at 4 years and 1 of each of the four at 5 and 6 had just entered when the second record was taken.

It was entirely in English, with an error index of only .040. When the second record was taken, his Chinese grandmother was present, and he tried to talk to her in Chinese. His knowledge of that language was much less than his knowledge of English: so his per cent of English words and sentences decreased and that of mixed sentences increased markedly.

This case illustrates the inadequacy of a single sample as a measure of an individual's proficiency. Such brief samples can be used only when collected in large numbers, whether for group or individual comparison.

The mixed sentences show a very erratic progress due to the heterogeneity of the subjects, 13 of whom were Chinese, 9 Korean 14 Filipino, 6 Japanese, and 2 Hawaiian. At the time of the first record, they varied all the way from using 100 per cent English words to only 29 per cent.

L. Is the Handicap Due Only to Pidgin English or Also to Bilingualism?

From the previous discussion, it is apparent that the non-haole children in Hawaii are retarded in language development. Not a single racial group studied has attained at six years the use of sentences as long on the average as those that the five-year-old Caucasian monoglots use. When they speak English, the number of errors per thousand words averages for every group higher than the number which the monoglots make at three years of age.

The question arises as to whether this retardation is sufficiently explained by the pidgin English current in the territory or whether the bilingualism of the majority is also a handicap.

Since many of the groups use very little but English, it would seem that the pidgin English is the main handicap; but when we compare the groups according to degree of bilingualism, the picture is different.

The homes of the Portuguese and Hawaiians both received language ratings of 3.0, which means that most of the children hear little but English in their homes, although that English is mostly pidgin. The Chinese and Korean homes both received ratings of 2.5. Some of them heard only English, but almost all heard a great deal of Chinese or Korean. The Filipinos and Japanese had the lowest ratings. The Filipinos heard much pidgin English and some Filipino, the Japanese some pidgin English and more Japanese.

Throwing the racial groups into three categories according to home language rating and comparing their average scores on the most significant criteria, we find, as shown in Table 44, that the groups whose homes' language rating is highest not only make fewest errors, inflect more verbs, and use more connectives than do those rating lower, as would be expected when English only is considered, but that they also use longer sentences and fewer interjections even when other languages are considered.

TABLE 44

Comparison of Groups According to the Average Language Ratings of the Groups as a Whole

Groups the language rating of whose homes is	3.0	2.5	2.2 or less
Average error index including errors of mixture	402	442	520
Average error index excluding errors of mixture	390	410	472
Average sentence length	3.20	3.15	3.03
Average per cent of English words that were interjections	2.6	5.0	9.7
Average per cent of English words that were connectives	3.8	3.4	2.9
Average per cent of verbs inflected	19.2	16.2	11.3
Average per cent of non-English words that were interjections	1.9	3.0	5.6

To eliminate the factor of a possible race difference, the two Filipino groups may be compared. The rural is more bilingual than the city group, whether the proportion of the different languages used, the language rating of the homes, or the number of children using more than one language is used as the criterion. The city Filipinos excel the rural on every measure of language development.

There are very few children observed from homes where pidgin English was not spoken. Homes where only good English was used by members of the household received a rating of 5.0; those where the English used was good but where another language was also used received a rating of 4.0.

It was possible to find 14 pairs—one member of each from a home rated five, the other from a home rated four—who differed not more than three months in age, or not more than one month in the cases of pairs less than three years of age; whose parents had received on the average within a half-year of the same amount of English education; and whose fathers' occupations on the Barr rating scale averaged less than one point. The pairs were of the same sex except for two pairs of four and six years, at which age the sex differences are very slight, and the boys are in the group which the differences in parental education and occupation favor.

TABLE 45

COMPARISON OF CHILDREN FROM HOMES WHERE GOOD ENGLISH ONLY IS SPOKEN WITH THOSE FROM HOMES WHERE GOOD ENGLISH AND ONE OTHER LANGUAGE IS SPOKEN

	Home language rating	Parental education	No. of children		Age in months	Sentence length	Words	Per cent of		Error index
			Boys	Girls				English sentences	Mixed sentences	
Japanese	4	12.0	1	0	46.0	3.6	96.0	90.0	8	175
	5	12.0	1	0	45.0	4.7	99.0	100.0	0	138
Hawaiian	4	8.0	11	13	48.5	3.16	97.2	92.4	4.5	459†
	5	10.1	8	14	49.0	3.24	97.2	95.6	2.8	319
Portuguese	4	7.1	5	5	45.3	3.36	98.3	95.1	2.0	291
	5	10.1	3	1	47.5	3.37	100.0	100.0	0.0	303†
Chinese	4	12.0	5	0	43.0	3.13	84.7	80.0	16.7	385
	5	12.0	3	0	44.3	4.20	95.7	89.3	6.7	297
*Matched pairs**	4	8.9	6	8	44.0	2.6	93.6	89.9	5.2	499
	5	9.4	8	6	44.1	3.8	97.5	94.0	1.7	296

*In the case of 2 pairs the other language was Portuguese; of two other Chinese, of one pair Japanese and of the rest Hawaiian.
†One excessively high two-year-old score.

Table 45 (see line labelled "matched pairs") shows that the children from the homes using English have a marked advantage not only in the measures of facility in English usage but also in sentence length. Their sentences average 1.2 words longer, and they make only about 60 per cent as many errors, use only one-third as many mixed sentences, and use about four per cent more English.

If all cases from homes rated four and five are included in the comparison except those from an age level at which there is no child of the same sex from homes of the other category to match those represented, the results are as shown in Table 45. For this comparison, the means of age-level averages are used, and in no other way was an attempt at matching made. For each racial group except the Portuguese, the average of mixed sentences and amount of error is less and of English and sentence length, greater, for the group from homes speaking only one language. The Portuguese from homes rated five do use more English, fewer mixed sentences; but their sentences are almost the same length and the amount of error is a very little greater than for those from homes rated four.

Children from homes rated one and two were also compared, since in homes rated one practically no English was reported spoken; and in those rated two, pidgin English as well as a foreign language was spoken (see Table 46). Our measures, being mostly of facility in English usage, are not so adequate for this comparison. Although it was reported that no English was spoken in the homes rated one, this was true only in the case of the adult members, since the children used quite a little English learned, probably, from their playmates.

It was difficult to pair the children in the groups on factors other than race, sex, and age. Many were eliminated until only 37 pairs were left. The other criteria used were Barr rating, parental education, and birthplace. For many of the Japanese, data were available as to education in Japanese schools, and for their children, Japanese education was the educational factor on which they were matched. As to parental birthplace, the children paired must have both parents born in the United States or members of the same pair might have one or none born abroad. The home advantage lies with the children of homes rated two in all respects where a difference exists, except for the Japanese, on the Barr rating.

Nevertheless, the children from the monolingual homes in each racial group used longer sentences, the difference, however, being

TABLE 46

COMPARISONS OF MATCHED PAIRS OF CHILDREN FROM HOMES RATED 1 OR 2 ON LANGUAGE AS TO PROGRESS IN SPEECH

	Parents' years of education in		Per cent born in U. S.	Bar rating	Number of		Age in in mos.	Sentence length	Per cent of sentences in		Error index
	Japanese	English			Boys	Girls			English	Mixed	
Language rating 1											
Japanese	7.0	0.8	20	8.58	6	9	40.8	2.33	34.5	28.3	660
Chinese English education 50% primary			46	8.19	8	6	47.1	3.26	42.8	6.6	417
Rural Filipino 1.1 years			0	4.91	5	3	47.8	3.16	20.0	27.0	607
All children:											
below 36 mos. in age				7.34	7	5	27.3	1.79	32.2	16.3	793
from 36 mos. to 60 mos. in age				7.66	9	7	47.0	3.21	25.3	23.5	430
from 61 mos. up				7.77	5	6	65.3	3.67	36.7	28.0	386
Language rating 2											
	Years of education Japanese	English									
Japanese	6.9	2.3	30	7.98	6	9	40.7	1.99	31.9	21.2	642
Chinese English education 36% primary			57	8.21	8	5	47.1	3.21	48.4	4.7	399
Rural Filipino 1.1 years			0	4.91	5	3	47.8	2.88	32.8	18.0	560
All children:											
below 36 mos. in age				7.55	7	5	27.2	1.59	42.5	11.5	687
from 36 mos. to 60 mos. in age				7.51	9	7	46.9	3.03	68.8	14.0	462
from 61 mos. up				7.96	5	6	65.4	3.34	59.1	19.1	465

very slight in the case of the Chinese. They also did not make many more errors in the English they did use, although, as was to be expected, they used less. The Japanese children from monolingual homes, however, even used a trifle more English.

In order to discover at what age the child from non-English speaking homes had the advantage, the children of all racial antecedents were thrown together and then separated in three groups according to age: children who had not attained their third birthdays, those from three to five years, and those who had passed their fifth birthdays.

The children from homes rated one in each age group used longer sentences and fewer English sentences, although the oldest group used almost as much English as those from homes rated two. They used more mixed sentences also. For both groups of children, the number of these sentences increased with age, so it is presumed that they had not reached the point of descent in the curve. As to errors, the youngest children from non-English speaking homes made more, the middle group a little less, and the oldest group about four-fifths as many as the homes where pidgin English is spoken.

The groups are too small to draw any conclusions, but two points are suggested:

It would appear that a more correct English would be used by young children if those parents whose English is inadequate would use only their mother tongue when talking to their children and let them learn English from other sources, preferably kindergarten.

The advantage in sentence length found where only one language is used by the adults in the home suggests that the bilingual handicap is lessened when the sources of the two languages are different.

These comparisons point rather strongly to attempted bilingualism, as well as the use of pidgin English, as a source of retardation in speech; but these comparisons are according to home language used by the child himself.

The order of the groups as to amount of English used (see Table 10) is Portuguese, Hawaiian, Korean, city Filipino, all of whom use over 90 per cent of English words; Chinese and rural Filipino, who use 80 and 79 per cent respectively; and Japanese who use 50 per cent. The extent of English used indicates the degree of bilingualism, as in no case is the other language used more than half the time. The number of children in each group who use sentences entirely of a non-English language is another criterion of extent of bilingualism, and, as seen in *Table 17*, the order of groups is very nearly the same on this criterion as it is for proportion of English.

In fewness of errors, there are only two changes of rank. The Chinese move up to second place and the Japanese to sixth. These two groups have been found in several investigations (25) to make higher scores on intelligence tests of different types than do the Hawaiians and Filipinos, whom they supplant by the criterion of errors. In length of sentence, the city Filipinos use only seven per cent fewer English words than do the Portuguese, so that this displacement of rank is not very significant.

TABLE 47

COMPARISON OF CHINESE, JAPANESE AND FILIPINO CHILDREN WHO USED ENGLISH ALMOST ENTIRELY PAIRED WITH THOSE WHO USED TWO LANGUAGES DURING THE PERIOD OF OBSERVATION

	Number of pairs Chinese	Japanese	Filipino	All cases
All	21	9	10	40
Boys	13	6	7	26
Girls	8	3	3	14
Children using 95 to 100% English				
Parental education	7.2	8.1	3.0	6.4
Barr rating	9.97	9.57	5.24	8.70
Age in months	49.1	47.0	45.8	47.8
Per cent English words	98.6	98.2	96.2	97.9
Sentence length	3.64	3.03	3.45	3.46
Error index	288	360	469	350
Children using 27 to 89% English				
Parental education	6.8	7.7	3.0	6.0
Barr rating	9.91	9.69	5.32	8.47
Age in months	49.1	46.6	45.7	47.7
Per cent English words	58.0	51.4	59.2	56.8
Sentence length	3.26	2.88	2.86	3.07
Error index	386	577	616	486

To compare individuals rather than groups, children using over 95 per cent English while under observation were paired with those using from 27 to 89 per cent English, according to race, sex, age, parental education, and occupation of fathers. The results indicate, as shown in Table 47, that the group confining itself mainly to English makes many fewer errors in English and uses longer sentences than does that attempting to speak two languages.

There were no cases using less than 10 per cent English, so a comparison with those using only another language could not be made.

Although this evidence is insufficient to prove anything, still it all suggests that an important factor in the retardation in speech found in the preschool population of Hawaii is the attempt to make use of two languages.

M. TYPE OF WORDS IN MOTHER TONGUE THAT LINGER LONGEST WHEN IT IS SUPPLANTED BY ANOTHER LANGUAGE

Three of the groups studied used very small amounts of their ancestral languages. This gave an opportunity to study the types of words that linger longest when one language is being supplanted by another.

Miss Lee (16) had made such a study of Hawaiian, and her results, reworked for comparison with the Portuguese and Korean groups, are included. The results of this classification of words by interests are shown in comparison with that of Boyd's (5) on total words used by his child at four years, and the classification by different words by Brandenburg (6) of his child at four years and by Nice (23) of her eldest daughter. In the last case, the results found at different ages from eighteen months to four years were averaged (see Table 48).

The classifications of words are, to some extent, overlapping. Proper nouns are not included in our comparisons, but they were by Boyd. Some reclassifications and recalculations were necessary to make the different lists comparable.

Nouns referring to people were subdivided into two classes: those naming members of the family by common nouns such as "mother" or "aunt"; and those referring to other people. The average per cent of total words constituted by such words for the three groups is almost 30, and they compose about 11 per cent of the different words of the three languages used. These are much higher percentages than those found for the English words. The percentage of such words is highest in the case of the two languages spoken only by descendants of that linguistic stock. Hawaiian, it will be remembered, has come to be used to a small extent by all the groups.

Only 106 Portuguese words that were not proper nouns were used. Over two-fifths of these referred to people, three baby words, "*vo vo,*" "*vôvô,*" and "*titi,*" referring to "*grandmother,*" "*grandfather,*" and "*aunt,*" respectively, comprising almost all such words.

Another group of words which compose a larger proportion of the mother-tongue vocabulary is that referring to food and eating.

TABLE 48

WORDS OF ANCESTRAL LANGUAGE RETAINED CLASSIFIED ACCORDING TO INTERESTS AND COMPARED WITH SUCH CLASSIFICATIONS AS MADE BY NICE, BRANDENBURG AND BOYD

	Hawaiian Per cent of 111 different words	Hawaiian Per cent of 562 total words	Portuguese Per cent of 58 different words	Portuguese Per cent of 106 total words	Korean Per cent of 104 different words	Korean Per cent of 338 total words	Average per cent of different words	Average per cent of total words	Boyd's per cent of total words at 4 years	Nice's per cent of different words at ages 1½-6	Branderburg's average per cent of different words at 3 and 4 years
Family	1.8	9.1	5.2	35.3	7.9	27.3	4.9	24.1	3.0*	2.5	6.5
People	7.2	11.4	8.6	4.5	3.0	0.9	6.2	5.6			
Food	17.2	16.4	13.8	10.4	13.5	21.0	14.8	15.9	0.9	4.8	7.0
Body, parts of	4.5	0.9	3.4	3.8	5.9	2.4	4.6	2.3	1.0	3.0	2.3
Religion	2.7	0.6	1.7	0.9	0	0	1.5	0.5	—	—	—
Health	1.8	2.1	3.4	1.9	3.9	2.4	3.0	2.1	—	0.8	2.1
Clothing	2.7	0.5	3.4	1.9	5.9	5.3	4.0	2.6	1.0	3.1	3.1
Play and occupation	5.4	5.2	3.4	1.9	3.9	2.5	4.2	3.2	1.0	4.5	7.7
Work	4.5	1.8	0	0	0	0					
House and furnishing	5.4	1.3	3.4	1.9	6.8	2.8	5.2	2.0	0.7	6.8	9.5
Other's belongings	0	0	0	0	3.9	1.5†	—	—	—	1.2	—
Topography	3.6	0.9‡	0	0	0	0	—	—	11.5	2.5	3.1

TABLE 48 (*continued*)

	Hawaiian Per cent of 111 different words	Hawaiian Per cent of 562 total words	Portuguese Per cent of 58 different words	Portuguese Per cent of 106 total words	Korean Per cent of 104 different words	Korean Per cent of 338 total words	Average per cent of different words	Average per cent of total words	Boyd's per cent of total words at 4 years	Nice's per cent of different words at ages 1½-6	Branderburg's average per cent of different words at 3 and +years
Abstract time quantity and indefinite	1.8	2.9	1.7	0.9	0	0	1.2	1.3	14.6	0.9	8.4
Vulgarities	7.2	8.9	5.2	3.8	1.0	3.8	4.5	5.5			
Bodily needs	7.2	15.5	0	0	3.0	0.9	3.4	5.5	21.1	—	23.2
Other verbs**	2.7	5.5	10.3	5.7	20.4	10.9	11.1	22.1			
Animals	0.9	0.2	5.2	3.8	3.9	2.3	3.3	2.1	0.9	8.4	3.8
Flowers and trees	3.6	3.4	3.4	1.9	1.0	0.3	2.7	1.9	0.5	3.2	1.3
Pronouns	0.9	0.4	1.7	0.9	0	0	1.2	0.7	18.1	1.5	1.6
Adjectives of quality	6.3	1.7	2.6	1.9	1.0	0.6	3.3	1.4	4.4	12.1	11.0
others	0	0	4.5	1.9	1.9	4.1	3.1	2.0	10.2		
Adverbs yes and no	0	0	5.2	8.5	4.1	4.4	3.3	4.3	11.6	5.5	4.8
others	2.7	0.6	6.9	3.8	4.9	2.3	4.6	2.2			
Interrogative words	1.8	0.9	3.4	1.9	1.0	1.5	2.1	1.4	—	—	—
Interjections	2.7	3.9	1.7	0.9	1.0	1.5	1.8	2.1	0.3	1.2	0.7
Prepositions	0	0	1.7	0.9	1.0	0.6	0.9	0.5	5.7	1.1	0.9
"Pau"	0.9	17.4									

*Includes proper nouns.

†All referred to purse or money.

‡All referred to sea and ships.

**Mostly verbs of motion-slap, switch and spank among them for last two groups.

This is not surprising, as most of the eating of young children is done at home, and foods of other lands often have no English name. In the case of Hawaiian, some of the native Hawaiian foods have come to be used by non-Hawaiians, with a resulting Anglicization of words such as "*poi*."

Other types of nouns comprising a somewhat higher percentage in the mother tongue than would have been the case had it been the only language used are words referring to clothing, parts of the body, health; and of the total, but not different words, those concerning the house and its furnishings or equipment.

Nouns expressing religious ideas composed most of the abstract nouns in the vanishing languages. All abstract and indefinite nouns and those expressing quantity and time were much less frequent in these languages than in the English vocabularies of the three monoglot children. The only nouns classified under "topography" or "civilization" used in Hawaiian refer to the sea and boats.

More words in proportion referring to play and occupation are used in Hawaiian than in the other two non-English languages. In Hawaiian, they form a higher per cent of total words than such words did in Boyd's list; but they comprise a much lower per cent of different words than in the Brandenburg and Nice lists.

Animals were very few in Hawaii before the white man came. Not only do almost no animal names occur in the Hawaiian list, but such words are less frequent than nouns referring to flowers and trees in that language, whereas the opposite is true in Korean and Portuguese. Both these classes of words, however, occur less often in the last two lists named above than in the English lists.

In all three languages, interjections comprise a higher proportion, while verbs, pronouns, modifiers, and connectives comprise a lower proportion than in the English lists.

One word, "*pau*," comprises over a sixth of the total Hawaiian words. "*Pau*" may serve as several parts of speech and as used did not fit any one class, so it was listed separately. It may be translated by "*finish*," "*done*," "*enough*," and similar words.

Miss Lee (16) had considered the frequency of vulgar and impolite words. She found such Hawaiian words used relatively more commonly by the non-Hawaiians than by the Hawaiians, but at least the Portuguese and Koreans use such words in their own languages too, as is shown in the table.

To summarize: the parts of speech that linger longest in competition with a new language seem to be nouns and interjections. Among the nouns, those referring to the more intimate aspects of the home are last to disappear. Among the very last are words expressive of family relationships and names of foods.

N. COMMONEST ENGLISH WORDS

An important part of the present study is the comparison of the English vocabularies of the children observed with those of normal mainland children. Such a comparison should be useful to the schools in the Territory in determining where emphasis must be placed or explanation most needed in the early stages of reading and oral language. It is planned to expand this part of the study in considerable detail; but in this monograph, only the commonest words are studied.

For this comparison three previous studies are used: Smith's (30, p. 26) list of all words occurring over 100 times in a study of English-speaking monoglots in Iowa in which the same technique was used as in the present study; the International Kindergarten Union's study (14) directed by Madeline Horn of words used by American children at kindergarten and at home; and the study of reading vocabulary suitable for the first three grades as made by Gates (12).

The three lists differ as to their methods of recording. Rank of words in the Gates' list is determined not only by their frequency of occurrence in the spoken vocabulary of children and in children's literature but also on experts' judgment as to their interest, utility, and difficulty. Derivatives were not often rated separately and the words were classified under parts of speech so that the same word form might occur in different places with different ranks. The International Kindergarten Union's list (14) listed all inflections and contractions of a word separately but not the different meanings of the same word forms. It included children's words, slang and colloquialisms. The Smith (30, p. 26) list included no derivatives unless they had a different meaning than the original word; e.g., "*going*" in the phrase "*going to*" differs from "*go*" in that it is used as a form of the future tense rather than expressing the idea of motion. Word forms, the same in sound and appearance but different in meaning were separately classified.

It had been planned to list separately both derivatives and word forms of different meanings but with the variety of help used in

TABLE 49

One Hundred Commonest English Words Compared with the Commonest English Words According to the International Kindergarten Union, Gates and Smith Studies

Common English words	According to Hawaiian study		Rank of word in Smith list	International** Kindergarten Union's list of words spoken at		Rank of words in Gates study
	Frequency	Rank of* word		Kindergarten	Home	
a	270	86	7	–/	–/	30
ah	1,049	22		—	—	—
all	628	35	43	–/	–/	17
am	71			–/	–/	54
an	11			–/	—	63
and	575	39	26	–/	–/	174
are	55			–/	–/	34
at	112		56	–/	–/	20
away	332	78		—	—	107
baby	465	52	49	—	–/	84
be	95			–/	–/	16
because	40			–/	–/	393
bed	18			—	—	78
big	397	64	45	–/	–/	39
book	160			—	—	83
boy	516	46		—	—	26
broke	369	68		—	—	—
but	78			—	–/	206
buy	297	82		—	—	177
by	156			—	—	71
by-and-by	348	75				
can	870	27	23	–/	–/	130
candy	273	84		—	—	386
can't	177			–/	–/	—
car	357	71		—	—	228
catch	252	93		—	—	—
come	1,487	15	38	–/	–/	29
daddy	555	44		—	–/	256
day	29			—	—	37
did	35			–/	–/	65
didn't	22			—	–/	—
do	146		6	–/	–/	14
does	3			—	–/	151
dog	97			—	—	75
don't	427	58	—	–/	–/	415
down	890	26	35	–/	–/	
eat	445	53		—	—	50
eh	474	51				
eye	100					87
fall	253	92		—	—	153
father	69			—	—	68
fire	73			—	—	99
for	497	49	64	–/	–/	15
four	58			—	—	58

**If word is checked –/, it is included in the first hundred most frequent words; if checked — in the second to fifth hundred, if not checked, it occurs lower than the fifth hundred.

*Ranks of first hundred only given.

TABLE 49 (*continued*)

Common English words	According to Hawaiian study		Rank of word in Smith list	International** Kindergarten Union's list of words spoken at		Rank of words in Gates study
	Frequency	Rank of* word		Kindergarten	Home	
from	93			—	—	66
get	1,381	17	22	√	√	31
girl	172			—	—	57
give	1,330	18		—	√	43
go	4,554	3	15	√	√	19
going	2,369	8	36	√	√	309
good	336	37		—	√	187
got	179		24	√	√	192
had	51			√	√	42
hand	173			—	—	82
has	134			√	—	56
have	186		16	√	√	38
he	1,217	20	47	√	√	7
her	94		57	√	√	35
here	1,489	14	11	√	√	184
hey	574	40		—		
him	395	65		√	—	55
his	142			√	—	33
home	400	63		√	—	47
house	354	73		√	—	73
how	312	81		√	√	425
I	7,928	1	1	√	√	2
if	110			√	√	344
I'll	23			√	√	411
I'm	197			√	√	488
in	501	48	31	√	√	5
inside	325	79				
is	360	70	2	√	√	3
isn't	2			—	√	
it	263	88	3	√	√	22
it's	16			√	√	470
just	51			√	√	325
kind n)	696	33		—	—	311
know	761	31	61	√	√	499
let	202		34	—	—	172
let's	43			—	—	112
like	2,400	6		√	√	85
little	116		63	√	√	138
look	1,587	10	28	√	√	90
ma	495	50				
made	28			√	—	48
make	978	25	30	√	√	11
mama	1,572	11	55	—	√	414
man	324	80		—	—	28
may	27			—	—	72

TABLE 49 (*continued*)

Common English words	According to Hawaiian study		Rank of word in Smith list	International** Kindergarten Union's list of words spoken at		Rank of words in Gates study
	Frequency	Rank of* word		Kindergarten	Home	
me	3,286	5	17	–/	–/	23
milk	65			—	—	88
mine	350	74	46	–/	–/	
more	1,300	19		–/	—	
mother	193			–/	–/	44
my	1,570	12	13	–/	–/	25
name	203			—	—	100
new	40			—	—	70
no	4,209	4	25	–/	–/	adv. 9 adj. 10
not	610	37	9	–/	–/	74
now	649	34	32	–/	–/	104
of	58		60	–/	–/	24
off	60			–/	–/	adv. 134 prep. 133
oh	437	55	20	–/	–/	o 316 oh 279
O. K.	228	99				
old	45			—	—	52
on	573	41	19	–/	–/	8
one	1,534	13	27	–/	–/	no. 12 adj. 13
only	293	83		—	—	270
our	77			–/	—	27
out	251	94	53	–/	–/	adv. 81 prep. 80
over	560	43	67	–/	–/	110
play	830	28	54	–/	–/	36
push	220	100				
put	611	36	40	–/	–/	168
red	41			—	—	49
right	254	91		–/	–/	287
run	137			—	—	40
saw (v)	34			–/	—	64
say	71			–/	–/	76
school	269	87		—	—	125
see	1,047	23	18	–/	–/	18
she	259	89½	59	–/	–/	53
sit	186	62		—	—	126
six	34			—	—	93
so	57			–/	–/	240
some	392	66	41	–/	–/	
stay	402	62		—	—	
stop	78			—	—	67
take	711	32	51	–/	–/	317
tell	531	46		—	—	395
that	1,027	24	5	–/	–/	96
that's	423	60		–/	–/	97

TABLE 49 (*continued*)

Common English words	According to Hawaiian study Frequency	Rank of* word	Rank of word in Smith list	International** Kindergarten Union's list of words spoken at Kindergarten	Home	Rank of words in Gates study
the	2,324	7	10	✓	✓	1
them	1,478	16		✓	✓	61
then	227	97½		✓	✓	485
there	1,165	21	21	✓	✓	219
these	20			—	✓	adj. 201 pron. 200
they	113			✓	✓	51
think	111			—	✓	474
this	2,061	9	8	✓	✓	adj. 94 pron. 95
throw	242	95½				
time	443	54		—	✓	140
to	581	38	Prep. 58 inf. 12	✓	✓	4
too	431	57	conj. 48	✓	✓	194
train	42		50	—	—	190
tree	47			—	—	69
two	373	67		✓	✓	32
up	794	29	29	✓	✓	adv. 46 prep. 45
us	434	56		—	—	59
wait	259	89½		—	—	
want	365	69	14	✓	✓	154
was	103			✓	✓	77
water	242	95½		—	—	60
way (n)	187		65	✓	—	155
we	535	45	37	✓	✓	21
well	26			—	✓	
went	418	61		✓	—	136
what	764	30	39	✓	✓	91
when	172			✓	✓	127
where	566	42	44	✓	✓	220
who	206			—	—	62
why	230	97½		—	✓	222
will	17		33	✓	✓	276
wish	4			—	—	92
with	277	85	66	✓	✓	272
won't	21			—	✓	
work (v)	88			—	✓	79
would	5			—	✓	195
yeah	356	72				
yes	424	59	42	✓	✓	120
you	5,891	2	4	✓	✓	6
your	344	77	52	✓	✓	41
yours	127			✓	—	

tabulating, the latter aim was found impracticable. All words in each conversation had been tabulated separately for each child in rough alphabetical order, under the different parts of speech in order to make that analysis. These lists facilitated the tallying, both as to number of children using and the frequency of occurrence of each word, against a master list of the most common words found in the Kindergarten Union's list.

Table 49 includes all words occurring in the Smith list (67 in all) and those occurring among the first 100 words in any of the other four lists; viz., those of the present study and of Gates' reading vocabulary, and the kindergarten and home lists of the International Kindergarten Union Study. This makes a total of 173 different words.

In the table the frequency in the Hawaii list of all 173 is given but only the first 100 are ranked. All the words in the Smith list are checked as to rank. All words occurring in the first 100 in the two Kindergarten lists are checked in one way, those in the next 400 are indicated by another symbol. The ranks of all the words on the Gates' list are shown if they occurred among the first 500. No word, however is included if in one of the five lists it was not ranked among the first 100 words.

Wherein the Hawaii list differs markedly from the other spoken vocabularies, a difference as to English usage exists which, in many cases are incorrect usages. Wherein it differs from the Gates list in much lower usage of particular words, it indicates common words with which the majority of local children may be unfamiliar at the time of school entrance.

A displacement of rank in lists compared among the words in the first quartile, where the frequency in the Hawaii list runs close to or above 1000 and the comparative frequency between words differing by but a single rank by tens or hundreds, is more important than it is in the lower quartiles. Some of the differences are due to the different methods of recording.

Five interjections *"ah," "eh," "hey," "oh,"* and *"yeah"* occur in the Hawaii list, only one of which *"oh,"* occurs among the most frequent words in any of the other lists. Attention has been called above to the high percentage of interjections used by our subjects.

In the use of pronouns there is considerable difference in ranks. In Hawaii *"you," "them,"* (commonly pronounced without the *"th"*) *"us," "me"* and *"he"* are favored at the expense of *"your,"* *"they"* and *"it," "our," "mine," "his"* and *"she."* (see Table 2 in Appendix *C*). *"My," "I," "him"* and *"we"* show no displacement. Aside from the first person singular pronoun, possessive forms are rare and there is a tendency to make one form do for several. Masculine pronouns are learned first and *"'em"* is substituted for *"it."*

Errors in the use of articles accounts for the comparatively low rank of *"a"* and *"an"* and the higher rank of their substitute, *"one"*. *"No"* is the fourth most frequent word in Hawaii owing to its incorrect use for *"not"* and failure to use the contraction of *"not"* with verbs; the only verb used with this contraction to any extent being *"don't"*. *"Yes"* has a somewhat lower rank, perhaps because the colloquialism *"O.K."* is used instead. The latter word occurred in no other list.

The Kindergarten list included several past verb forms among the *first* 100 words; the Hawaii list includes only two, *"broke"* (which word the children know only in that form) and *"went"* which is sometimes used as an auxiliary in forming the past tense. *"Go"* another word so used and also, with *"going"*, almost exclusively as an indication of the future as well as redundantly holds the third rank. Other than with these few words little attempt at conjugation occurs, as is evidenced in the extremely low frequency of other past forms and absence of the different forms of the auxiliaries *"may," "will"* and *"be."* The use of *"like"* for *"want," "make"* for *"do," "tell"* for *"say," "get"* for *"have,"* and *"stay"* for the copula *"is"* account for the higher rank on the Hawaii list of the first word and the lower rank of the second in each pair of words.

Most of the prepositions and conjunctions in the table are used relatively less frequently in Hawaii. One of the exceptions *"for,"* however, gains over the Smith list due to the tendency to use it instead of *"to"* with the infinitive. *"Too"* occurs with high frequency, not as a conjunction as in the Smith list, but as an adverb used not only correctly but also as a substitute for *"very"*.

"Ma" and *"mama,"* which are similar to or the same as baby words for *"mother"* in languages other than English spoken in Hawaii, rank 59 and 11 respectively, higher than in other lists but *"father"* and *"mother"* are rarely used. It will be recalled that words designating family relationship are the last to drop out of home languages which are competing with English.

The phrases *"waste time"* for *"don't want to"* and *"all time"* for *"always"* account for the high frequency of the word *"time"* in the Hawaii list and the use of *"kind"* similar to that of a Chinese classifier and in other redundant ways gives it a rank of 33 in that list although it does not attain the first 100 in a single mainland list.

"More" ranks 19 in the Hawaii list and places in the first 100 only in one of the Kindergarten lists. Its high frequency is due to its use as the sole means of comparison by the majority of our subjects and its occurrence in the phrase *"no more"* used for *"have none"*.

"Over" gains in rank from 67 in the Smith to 43 in the Hawaii list owing largely to its frequency use in the phrase *"over there"* (frequently slurred so as to sound almost like a single word *"o-'dere"*) where mainland children would use the word *"there"* alone.

The other words used relatively more frequently in Hawaii are probably so used, either because some word must take the place of the displaced words or because of some peculiarity of the situation. For example the higher frequency of *"man"* and *"candy"* is found to be due mainly to the Filipino children whom the observer, a *man*, was accustomed to treat to *candy* in order to gain rapport.

Common words in the Gates' list found with very low frequency in the Hawaii list are *"bed," "day," "dog," "eye," "fire," "hand," "milk," "train," "tree," "off," "out," "new," "old," "red," "six," "run," "sit," "stop,"* and *"work."*

Since *beds* are articles of furniture often absent in Japanese homes, the habit of drinking *milk* has to be taught the children of oriental immigrants, *trains* are very few in Hawaii so that many of the children had not seen any, and *fires* are not needed for warmth in Hawaii, the rarity of such words may be due to the child's environment and not his inadequate English vocabulary. However, when to the score of words listed above are added the varieties of verb forms, especially of the verb "to be," the prepositions and conjunctions used but rarely and by many of the children not at all, it is evident that there is a serious lack of vocabulary on the part of a large proportion of school entrants in Hawaii sufficient to handicap them to a considerable extent in learning to read.

V. SUMMARY AND CONCLUSIONS

A. SUMMARY

1. A study has been made of the language development of children from two to six years of age, born in Hawaii, who come from a non-English-speaking ancestry.

2. Six samplings of 125 each of children residing in Honolulu and representing groups of Chinese, Filipino, Hawaiian and part-Hawaiian, Japanese, Korean, and Portuguese ancestry have been studied. Besides these, records for 125 Filipino children residing in rural areas were gathered, and data received on sixteen Japanese and four Hawaiian children living in the country were examined. There are also included two groups totalling 125 children of mixed racial antecedents who were studied in other than the standard situation. These samples are fairly representative of Island groups except the Korean, who are somewhat selected.

3. The method used was the verbatim recording of each child's spontaneous conversation by a student who could speak the language of the child's ancestors in addition to English.

4. The standard situation in which the record was made for the first seven groups mentioned above was in or about the child's home while he was at play with his siblings or other playmates. Two other situations were used for smaller samplings: 100 children were studied during free play periods at kindergarten or nursery school, and 25 were studied in their own homes with only adults present.

5. The conversations so obtained were analyzed as to proportion of English used; average number of words per sentence; degree of egocentricity shown; types of sentence classified according to form and function; number of negative sentences; number, function, and form of questions; parts of speech; and inflection of words. The words most commonly used were also determined for English and for a few of the other languages used.

6. The results of this analysis were studied to determine the effect of race, sex, parental background, home influence, order of birth, residence, kindergarten attendance, and the different situations in which the records were taken.

7. The results show that in comparison with haole children and those studied on the mainland, the Island children are seriously retarded in the use of the English language, a retardation which is not compensated for by a greater advancement in other languages used.

8. On the whole, the children seem to prefer to use English if they can, even when another language is spoken in the home. Even the Japanese children, who use it least, use about 50 per cent English. About 88 per cent of the words recorded in the study were English, or 87 per cent if only the seven main groups are considered. These groups use 12 per cent of words from home languages other than English and about one per cent of words, mainly Hawaiian, that filtered in from other languages spoken in the Territory.

9. In Honolulu the children of Portuguese ancestry hear and use little but English; the Hawaiians hear some Hawaiian and the Koreans, considerable Korean, but each group uses 96 per cent English words; the Filipinos hear some Filipino and much pidgin English and use 92 per cent English words; the Chinese use 80 and the Japanese 50 per cent English words.

10. The average per cent of sentences entirely English is less for each group than the proportion of English words, since many sentences composed of more than one language are used.

11. Except for the Portuguese, who use very little but English, the per cent of English, whether of words or sentences, increases with age up to six years.

12. Mixed sentences, that is, those composed of more than one language, tend at first to increase in number with age and then to decrease for the more bilingual groups. For the Hawaiians they decrease and for the Portuguese increase in frequency with age. Such sentences are especially frequent among the Japanese and Filipinos; but whereas such sentences are for the former almost always a combination of English with the parents' mother tongue, for the Filipinos there are many sentences composed of three or four languages or of two languages other than a Filipino dialect.

13. The average amount of error in the use of English by the different groups varies from 379 per thousand English words for

the Portuguese to 587 per thousand for the rural Filipinos. The order of the groups in correctness of usage is as follows: Portuguese, Chinese, Hawaiian, Korean, Filipino, and Japanese. Over half of the differences found in comparing each group with the others is significant. The average error indices are all much higher than that of monolingual Caucasian three-year-olds as found in a previous study (28).

14. The frequency of error tends to decrease with age, but the only marked decrease for most groups is from two to three years. On the other hand, the two most monolingual made more errors at six than at five years.

15. The most frequent type of error found is the use of incomplete sentences, and, next, the incorrect use of verbs. The insertion of words from other languages and the omission or incorrect use of prepositions and infinitives come next. The error type showing greatest excess over that found for Caucasians in the former study is in the use of the negative.

16. There is some difference between groups as to pattern of error and in the frequency of certain peculiar error idioms, which in some, but not all, cases can be traced to the influence of parental language. For example, a common error among the Japanese is the use of *"me, I"* as the subject of a sentence. Such repetition of two forms of the pronoun is good Japanese.

17. Error types that decrease in frequency with age consist of incomplete and mixed sentences and the wrong use of articles. No other types show a consistent decrease; in fact, in some cases an increase is found which, however, is at least partially explainable on the ground of the greater opportunity for making such errors when more complex sentences are attempted.

18. The various groups use, on the average, shorter sentences than do children in a less polylingual environment. The average length is slightly more than three words, which is about the average found in other studies at three years.

19. The rank order of the several groups from longest to shortest sentences used is Filipino, Portuguese, Chinese, Hawaiian, Korean, and Japanese. Eleven of the 27 differences between the groups are significant. The sentences increase in length with the children's increase in age. The differences from two to six years or from two

to three years are all statistically significant, and also those from three to five years in the case of four groups.

20. The longest sentences used are those composed of words from more than one language; next are those entirely in English. Except for the Chinese and Japanese groups, very few sentences of more than one word which are composed entirely in a language other than English are used.

21. The degree of egocentricity, or tendency to use sentences with one's self as the subject, as determined by Fisher's (11) method, shows great individual variation. It is lowest at two years and is fairly constant thereafter. The Filipino and Portuguese have the lowest average index; the Hawaiians have the highest. The more bilingual groups use more sentences that are classified as non-verbal than do the others.

22. The children in Hawaii were compared with a monolingual white American group previously studied (31). They are found to use more exclamatory and slightly fewer interrogative sentences, and to make much less frequent use of complex and compound sentences. Sentences that serve merely to name an object or person continue to a later age than with monoglot children. However, age trends are found to be similar, for exclamatory and naming sentences decrease; questions, answers, and complex and compound sentences increase, with age. Racial differences are not marked in the type of sentences used.

23. Questions asked by the children in the different groups number from 479 asked by the rural Filipinos to 987 by the Hawaiians. The city Filipinos asked 573, the Portuguese 528, and the Asiatic groups from 633 to 694 questions. From 25 to 73 per cent of the questions were directed to adults rather than to playmates.

24. Fewer questions of causality are asked than in the case of Caucasian monoglots, and more, especially among the more bilingual groups, are asked inquiring for the names of objects or as to what had been said to them. These differences indicate less maturity in speech. The monolingual Portuguese are an exception.

25. Questions in English are usually formed either by the use of an English interrogative word or by the addition of an interrogative particle such as *"eh?"* or even by adding a similar Japanese or Chinese particle. The Koreans and a few of the other children

make use of the phrase *"you know"* instead of *"isn't it"* in questions seeking corroboration. Interrogative words increase in number with age, whereas their use decreased with age in the case of the Caucasian monoglots.

26. The analysis made of parts of speech shows that in the use of connectives and pronouns and in the reduction with age in the proportion of interjections used, the Island groups are retarded, as they are in the use of the copula and inflected forms of verbs and nouns.

27. Much less use is made of articles, and of those used, the definite article occurs almost exclusively.

28. There is a slight superiority of girls over boys according to the various criteria of mastery of English used.

29. Children whose parents were born in the United States use more English and use it somewhat more correctly than do those with one or more parent born elsewhere; but there is no consistent difference between these groups as to length of sentence.

30. Children whose fathers' occupations rate above 9.0 on the Barr scale tend to use more and better English and longer sentences than do those whose fathers' occupations rate below 9.0. The difference is, however, not great for most groups, and in the case of the Koreans, does not exist as to proportion of English.

31. In the Hawaiian group, the part-Hawaiian children who represent more than two racial stocks have a slight advantage in mastery of English, according to the various criteria used. Then, in order, come the Caucasian-Hawaiian, Asiatic-Hawaiian, Filipino-Hawaiian, and, last of all, the pure Hawaiian.

32. By all criteria of mastery of English, there is a positive relation with the number of years of English schooling of the parents. The degrees of relationship differs for different groups and with the criterion used.

33. The most marked relationship found between mastery of English and possible factors studied is with the language rating of the home: children from homes where more and better English is used by other members of the household speak more and better English themselves and, to a slight extent, use longer sentences.

34. More English is used in the homes of Filipinos where the parents speak different Filipino dialects than where they speak the

same. The children whose parents use different dialects tend to speak more correct English and to use fewer mixed sentences, although their sentences tend to be somewhat shorter.

35. Later-born children in a family where a foreign language is spoken tend to use more English than do earlier-born children of the same age; but in most groups there is very little difference between older and younger children as to amount of error. In five of the seven comparisons made, slightly longer sentences are used by the earlier-born children.

36. Younger children tend to surpass their elder siblings in reaching the racial norms for their age in the use of English words but not in absence of error or in the length of sentences. The older siblings, except for the Filipinos, use longer sentences, even when allowance is made for difference in age, than do the younger.

37. More English is used and it is used more correctly in Honolulu than in the rural areas. The city children also make greater use of the conjugated forms of verbs. The Filipinos in the city use longer sentences also. More Hawaiian is spoken in the country by each of the three races observed.

38. Children who reside in neighborhoods where most of the residents are of the same racial antecedents as their parents do not make so much progress in learning English or in its correct usage as do those children who reside in mixed areas.

39. Records taken at kindergarten show that much more English is spoken there than about home. Longer sentences are also used and the language is more correctly spoken.

40. Bilingual children who attend kindergarten or nursery school gain so much practice in speaking a more nearly correct English while at school that there is a carry-over into the home.

41. In a situation with adults only present, the children use more English, longer sentences, and make fewer errors than they do when playing with other children.

42. Children who were studied more than once, the second time after intervals averaging about nine months in length, show increase in use of English, reduction in errors, and increase in sentence length. This agrees with the trends found at different age levels in the several groups.

43. An attempt made to separate the influence of the two factors,

bilingualism and pidgin English, on the retardation of speech of the children studied in Hawaii suggests rather strongly that not only is the child's mastery of standard English hindered by the prevalence of incorrect pidgin English, but also by his attempt to learn two languages.

44. When a language is disappearing before competition with a new language, nouns and interjections linger longest. Those words that refer to the more intimate aspects of the home, especially words expressive of family relationship and pertaining to eating, are the last to be abandoned, at least by young children.

45. The commonest English words, when contrasted with those found to be most common in mainland studies, give further evidence as to the high frequency of interjections; the failure to inflect words or to use the copula or indefinite article; and the tendency to substitute wrong verbs for the correct ones and the words *"no"* and *"o.k."* for *"not"* and *"yes"* and to confine the use of pronouns to only a few forms. The absence or rarity of words found by Gates to be among the most essential in a child's reading vocabulary suggests a severe handicap in learning to read, due to lack of vocabulary on the part of school entrants in Hawaii. The nine most frequent words are *"I," "you," "go," "no," "me," "like," "the," "going,"* and *"this,"* each occurring more than two thousand times.

B. CONCLUSIONS

The children in Hawaii from non-haole homes are retarded in language development to a degree so marked that, on most criteria, at the time of school entrance they are at about the level of three-year-old children from a less polyglot environment. The retardation is due to two handicaps: the prevalent use of pidgin English, and the bilingualism of many homes.

Children have a marked advantage over other children in mastering English when their parents were born in the Islands, received at least grade school education, have higher socio-economic status, and speak only English or, if they speak another language, at least use good English in the home.

When the mother tongues of the parents differ, the tendency is to use more English in the home, which condition leads to an earlier mastery of English by the child.

Kindergarten or nursery school attendance is a definite advantage for the bilingual child in Hawaii, particularly if there are more than a few Caucasian children in the same school.

It would be desirable, therefore, for such schools to be added to the local school system. There are, however, suggestions of slight contamination of the speech of children who speak better English, so that the number who use poorer English should not constitute too large a majority in any one school if such attendance is to improve the children's speech.

The tendency of racial groups to congregate in separate residential areas, while it may be desirable for some reasons, is not conducive to the preschool child's advancement in the use of English.

Older brothers and sisters help the younger child to use more but, unfortunately, not better English.

The extent and correctness of English spoken by children in each racial group are related to the date of their ancestors' first arrival in Hawaii and their first contact with English, to the relative proportion of the population made up by each group, and perhaps to some degree to the kinship to English of the language spoken by the group.

In general, by far the greatest number of the preschool population prefers to use English rather than any other language. If they could but hear good examples, the coming generation would be quite Americanized in speech.

The evidence, although insufficient, suggests that pidgin English is more responsible for incorrect English and bilingualism for the overuse of interjections, short sentences, immature type of questions when classed as to meaning, and lack of complex sentences.

APPENDICES

APPENDIX A:

FORM USED IN COLLECTING DATA

1. Date of Observation
2. Child observed
3. Date of birth
4. Age
5. Father's occupation
6. Mother's occupation
7. Father's education
8. Mother's education
9. Father's racial descent and birthplace
10. Mother's racial descent and birthplace
11. Other adults in the home
12. Does the child go to kindergarten?
13. Place
14. Residence
15. Sex
16. Order of birth
17. No. of children younger
18. No. of people present
    (adults
    (children
19. Language used by each

	preferred lang.	*2nd lang.*	*3rd lang.*	*4th lang.*
a. Child	....	....	....	....
b. Mother	....	....	....	....
c. Father	....	....	....	....
d. Grandparents	....	....	....	....
e. Other adults	....	....	....	....
f. Siblings	....	....	....	....

20. Time spent to take record
    Recorded by

APPENDIX B:

SAMPLES FROM CONVERSATIONS OF AVERAGE FOUR-YEAR-OLDS OF DIFFERENT RACIAL ANTECEDENTS

1. *Boy, 3 Years 9 Months, Parents Japanese, Both Parents Born in Japan, Some Japanese Schooling*

Oka San, this kind.	Going up on a box for something from his parents' store.
Pants no naka agate ga alu. (There are agates inside my trouser pockets.)	Mother gets some candy for him.
Pants no naka ile te kule nai. (Please put it in my trouser pocket.)	Mother gets some candy for him.
Shut up.	A friend of his mother asks him for money for the candy.
Two candy.	Goes out and gives two pieces to his brother.
Come on (come on)	Sees a group of boys playing marbles in an argument.
No go catch.	Sees a group of boys playing marbles in an argument.
Kakelo ya (Let's run)	Runs off himself.
Kule do this way (am coming)	Runs off himself.

Shut up, you.	Running around and playing with Akira.
Akira!	Running around and playing with Akira.
Ano, girl. (that)	Referring to writer.
Choose se (do)	
You nala kulai ban made olu? (You stay until dark night.)	Addressing the writer.
Wait.	Playing "hide and seek".
Safe.	Playing "hide and seek".
Utta (hit)	Hits friend.
Sure.	Imitating a boy.
Shut up.	Imitating a boy.

2. Boy, 4 Years Old; Parents Chinese; Father—Born in China, High School Education in China; Mother—Born in Honolulu, High School Education in Hawaii

Come, Shirley, ngo-de (we) *play school*	Speaking to Shirley (seven year old cousin).
I know one fish song.	Telling supposed teacher.
I no can sleep.	Telling teacher during sleep period.
Soff (soft)	When teacher spanked him for not sleeping.
I tal story. (tell)	Tell story.
I like play yo-yo.	Addressing another boy playing yo-yo.
This yo-yo not for him.	Addressing brother holding his yo-yo.
How can for him?	Speaking to self.
No more name.	Speaking to self.
For me?	Speaking to self.
Look! mama (Look! Look!)	Calling mother's attention while playing.
I can!	Speaking to self.
Good fun!	Speaking to self.
Mama, ngo oi (I like) *water.*	Speaking to mother.
Look! gum (so) *full go*	Refers to coffee on kitchen table.
Coffee dit (drops) *chut* (out)	Speaking to kid brother and mother.
Ngo (I) *put powder.*	Speaking to kid brother and mother.
Just like girl.	Speaking to kid brother and mother.
Mama, I like go riding wit Me-gao-jay. (aunty)	Speaking to mother.
Be ngo sarm juk. (give me clothes—wear.)	Speaking to mother.

3. Girl, Korean, 3 years 11 months, Parents Korean; Father Born in Korea, No English Education; Mother Born in Hawaii, Sophomore in High School

I no like only sing.	(Is asked to sing.)
I like sit down and sing.	
I no like sing.	"I don't want to sing." Cries.
Look Onnie.	Doesn't want to sing because Onnie is looking at her.

Hal-moo-nie (Grandmother)	Is asked who gave her the piece of meat.
My house get plenty, boy!	"We have plenty of candy at home."
I goin drink water.	
'At's mine, you know.	"The marble is mine."
'At's for dem, for Mah-chong dem.	"That's for them, for Ma-chong folks."
You like 'um, eh?	Holds marbles in hands.
Da towel wet.	Touches towel.
Ma, try look. I stay still.	Stands for mother to look at her.
Two fella go smoke and him.	Watches and counts the number of people smoking.
Go 'way.	
Mama, when we go home, buy candy you know.	
I went find 'um.	"I found it." Mother asks where she got marble from.
I goin to poke 'um.	Pokes stick in ground.

4. *Girl, 3 Years 11 Months, Parents Filipino, Both Born in Philippines, 7th and 4th Grade Education*

Me mines one.	Sister asked whose candy was on the bench.
I go live o' dea'.	Went into a small room.
I no scare.	Sister said there was a ghost.
Look at dat.	Sister soiled her dress.
I gon upstairs.	
I calabosin.	Jailing. She blocked the sister.
I no like go up.	Stairs.
I fall down.	
I go down again.	She was afraid to fall.
Well, you not to go someplace.	Talking to the baby sister.
You pake. (Hawaiian term for Chinese)	Calling sister.
I no gon give you candy.	
Yaiyai my likod. (sore my back)	Asked the mother to scratch her back.
No can catch me.	Ran away from the baby sister.
Darling no can catch me.	
I go tago. (hide)	She hid back of the door.
You da ibe. (shrimp)	Called the sister shrimp.
You io. (shark)	
You ibe and io.	
Me big like da man.	Like the observer.

5. *Boy, 4 Years 1 Month Old, Parents Both Born in Hawaii; Father—Portuguese-Hawaiian, 8th Grade Education; Mother—Haole-Hawaiian, 6th Grade Education*

Mama Kaluhi coming home.	(when he saw sister walking down the street.)
Eddie too	(seeing his brother coming home from school)
Pau school	(telling mother school was over)

Somebody coming mama	(when he saw O coming into the yard)
Beatrice we go play	(asking sister)
Baby you like play?	(asking little brother)
Play marbles	(suggesting what they should play)
I like play	(asking sister)
I get one agate	(telling sister)
You shoot Beatrice	(playing marbles)
Mama I like one piece of bread	(asking mother)
I like now	(when mother said it wasn't lunch time yet)
Ah there Bobo	(seeing little playmate coming into his yard)
Come Bobo	(inviting shy playmate to come nearer)
Us go play Bobo	(asking playmate)
Mama look Beatrice hitting me	(calling to mother from outside)
I going tell mama	(threatening sister)
Kaluhi look Beatrice	(calling to big sister)
Emily!	(scolding baby sister)
Mama no	(telling mother he did not want to take a bath)
I no like auau (bathe)	(to mother)

6. *Boy, 4 Years 2 Months Old, Parents Portuguese, Born in Honolulu, 3rd and 4th Grade English Education*

Get one big hole over there.	Playing on sand.
Push 'em.	Playing on sand.
Eh you guys!	Calling friends.
Your mother calling.	Telling friend.
Go-you.	Calling attention to friend.
Ketch 'em.	As they see a crab.
Where?	Looking for crab.
Went inside the water.	As crab crawls into water.
Ishie, try come.	Calling his little friend.
I stay all wet.	As wave splashes on him.
Look the dog.	As dog is seen swimming.
He stay swimming.	As dog is seen swimming.
You can swim?	Asking friend.
I can lili bit.	Telling friend.
I scare go far.	Telling friend.
Over here got sand.	Calling friends.
Come.	Calling friends.
We go make one puka (hole).	Building a tower-like.
Go 'way.	As dog comes around.
Lend me that.	Asking friend to lend him a spade.

APPENDIX C

TABLE A

PER CENT EACH TYPE OF ERROR WAS OF TOTAL ERRORS

	Portuguese	Hawaiian	Filipino		Korean	Chinese	Japanese
			city	rural			
Omissions of							
verb	4.3	5.6	4.5	5.2	3.9	6.7	7.3
subject or expletive	5.2	5.0	4.9	4.4	5.1	6.3	4.9
object	6.1	3.9	5.2	5.9	4.9	5.7	5.1
copula	5.5	5.6	5.7	3.8	6.0	4.9	3.9
verb and subject	15.2	15.8	7.8	10.7	14.4	15.1	20.9
verb and object subject and object	0.7	0.8	0.8	1.4	1.0	0.8	0.3
Verbs							
wrong choice	2.8	1.7	3.3	3.5	3.3	1.5	0.6
past (present for past or wrongly formed)	3.4	3.5	2.9	2.8	3.7	4.2	0.7
auxiliary omitted*	12.9	17.0	11.8	9.6	13.5	12.6	5.7
future	4.0	4.0	5.0	6.4	3.4	3.1	1.8
past used for present	0.2	0.1	0.3	0.2	0.2	0.3	0.3
present used for other tenses	0.5	0.3	0.5	1.0	0.4	1.1	0.2
other errors	0.0	0.1	0.1	0.0	0.1	0.1	0.2
Substantives							
gender	0.2	0.1	0.6	0.4	0.5	0.2	0.5
number	1.2	1.7	2.5	2.0	2.4	1.8	1.5
case	2.5	4.0	4.8	3.5	3.0	2.3	10.4
Prepositions	4.9	4.6	4.6	4.5	4.2	3.7	2.3
Infinitive sign	5.2	7.4	4.5	3.2	6.2	5.8	2.8
Articles	7.3	4.7	5.1	3.8	6.8	5.0	4.6
Modifiers	2.4	1.3	1.3	2.0	1.5	0.9	0.5
Connectives	0.6	0.5	0.7	0.7	0.6	0.9	0.2
Negatives	6.3	4.5	6.3	6.6	5.8	6.3	4.5
Comparison	1.0	0.2	0.2	0.1	0.1	0.2	0.2
Agreement	1.9	1.4	1.5	0.8	1.4	2.0	0.5
Order	2.0	1.1	2.0	2.1	1.7	0.7	1.1
Confused parts of speech	0.5	0.3	0.5	0.3	0.8	0.8	0.8
Redundancies	0.9	0.4	3.9	3.0	2.0	1.4	1.0
Miscellaneous	0.0	0.3	0.0	0.1	0.0	0.0	0.5
Mixed-in foreign words	2.2	4.0	8.7	11.9	3.0	5.6	17.1

*Includes a very few cases where a wrong auxiliary was used.

TABLE B

DETAILED LIST OF ERRORS IN THE USE OF PRONOUNS MADE BY EACH OF THE MAJOR GROUPS

	Portuguese	Hawaiian	Filipino		Korean	Chinese	Japanese	Total
			city	rural				
In gender—he for it	1			8	3			12
masculine pronoun used for feminine	3	7	50	27	30	9	2	128
In number (?)								
'em for it	69	39	158	112	121	37	25	561
this for these				1				1
In case								
1st person								
me for I	47	39	166	92	67	41	324	776
me for my or mine	6	11	27	11	2	2	16	75
my for I or mine				2				2
mines	4	2	24	17	2	3		52
mine for my	4							4
I for me		1	3	5	5	4		18
us for me	23	87	30	42	20	34	14	250
us for our or ours			1	2				3
2nd person								
you for your or yours	16	60	160	63	31	11		287
yours for you					1			1
yours for your	1		1					2
your for yours	1							1
3rd person								
him for he	38	13	19	25	17	2		114
him for his	2		6	4	6			18
he for his		1	3	2	4			10
her for she	13	9	1		2			25
hers for her	2							2
she for her		1						1
them for they	3	6	2		2	97	1	14
them for their					1			1
this for their				1				1
who for what		1	1					2
who for whose				3				3
who for whom	1		1	1	2			5
Japanese possessive added to English pronoun							70	70
Total	234	277	599	418	316	143	452	2437

(?) This may be confusion of gender, "*em*" might come from him.

Objective case used for another	1563
Possessive case used for another	3
Possessive case formed as in nouns (mines, yours, hers)	56
Other wrong form of possessive	8
Subjective case used for another	29

TABLE C

TOTAL NUMBER OF ERRORS IN THE USE OF ENGLISH MADE BY EACH GROUP

	Portuguese	Hawaiian	Filipino		Korean	Japanese	Chinese	Kindergarten	Rural Japanese	In Situation *A*
			city	rural						
Omissions of:										
verb	323	394	392	421	304	301	381	280	36	38
subject or expletive	383	351	433	357	392	203	358	236	27	55
object	447	269	460	479	375	210	320	260	17	40
copula	402	408	499	307	466	161	279	256	17	47
verb and subject	1113	1106	683	864	1107	858	853	723	83	183
verb and object	28	30	47	40	34	4	10	7	6	
subject and object	24	25	26	75	46	9	35	10	1	1
Verbs										
wrong choice	202	123	293	284	257	26	87	146	5	25
past: present for past	62		77	116			236	101	2	22
past		241			291	30				
past: wrongly formed	185		181	106				36	2	6
auxiliary	947	1187	1043	777	1041	235	711	511	11	121
future	293	275	437	513	264	73	177	207	14	23
past used for present	11	8	23	19	15	14	14	3	2	5
present used for other tenses	40	19	45	79	31	10	63	48	1	11
other errors	1	6	9		9	8	8		7	
Substantives										
gender	4	8	56	36	39	2	9	9	0	1
number	101	121	222	162	186	63	101	65	5	15
case	181	274	427	279	228	429	130	194	96	18

TABLE C (*continued*)

	Portu-guese	Hawai-ian	Filipino city	Filipino rural	Korean	Jap-anese	Chi-nese	Kinder-garten	Rural Jap-anese	In Situ-ation A
Prepositions	360	318	404	360	323	96	211	228	16	53
Infinitive sign	378	520	398	262	477	117	329	180	7	70
Articles	555	328	445	312	523	191	282	392	30	45
Modifiers	178	95	114	162	117	20	51	47	6	13
Connectives	44	33	59	54	50	10	50	29	5	5
Negatives	451	510	553	529	444	184	358	265	17	33
Comparison	73	15	17	10	9	7	9	17	7	3
Agreement	139	101	130	67	109	25	113	54		36
Order	150	76	174	166	129	46	42	167	28	29
Confused parts of *speech*	40	18	47	28	58	54	46	31	0	12
Redundancies	69	28	345	242	152	43	78	213	6	26
Miscellaneous		20		7	1			2		
Mixed-in foreign words	159	279	767	959	233	705	314	115	191	64
Total English words used	20350	18827	19858	15062	18779	8861	16378	14370	1145	5130
Total errors	7321	6984	8806	8072	7710	4112	5655	4832	643	1000
Errors per 1000 words	360	371	443	536	411	464	345	336	562	269

Appendix D:

Sentences to Illustrate the Special List of Errors

already	The car roll this ball all go down already.
all same	This all same red.
as why	Us no more money to pay, as why. Only sometime I stay home, the baby, as why.
been	Who been take that? You been see the towel?
break—tear	Break the paper.
broke—break	Ma, broke up the string. Frances broke the flower.
bring—come	Bring bottle come.
by-and-by	*By and by mother scold.*
chance	Whose chance for deal the cards?
every time	You every time lick me. Every time I stay home.
find	I finding it but I don't see it.
for	All right for buy milk nickel? *You get fish for eat?*
funny kind	You talk Chinese funny kind.
get	Upstairs get shoes. Way up there get one girl.
go	I like go pull. I go have the same cards again. By *and by you go fall down. Fred go bite (did bite) my* nose.
house	Us house no more baby. You house get plenty candy?
kind	Ah no quitting kind. This is marble kind agate.
lazy	I lazy stay in the house. Ah I lazy play. Lend me look. I lend her hold um.
like	He no like come. That's why you no like play? I like *buy candy.*
little more	Little more us going. Little more *pau* school.
long time	Long time went go work.
make	Goofy how you make. I make easy. Rachel no make. No make like that.
me, I	Me, I know.
more big	I run more fast than Helen.
more bigger	You going teach us more better.
no	You no can hear? No make hard.
no more	Us no more car. O'dere no more license.
no need	No need come. No need *hemo* (take off) your clothes.
one	Look, I get one gun. Baby swallow one seed. One train go down.
open—turn on, unbutton	Open the water. Open my pants.
plenty	When plenty boat come in. I see plenty ladybug.
scare	No scare!
shame	I shame. He no like go; he shame.
sore	O-oh sore! My teeth sore. No make sore.
stay	She mother stay home. Paper no stay outside. Blackie stay going.
take	Take 'em come. Take me go with you fellas.

tell	You tell you bring home candy. You no tell *mahalo* (thank you).
the	Oh the near you deal.
too	Too good. (expressing delight over the church festival)
try	Try lend me look. Come on baby, try come. Try catch me.
waste time	Waste time play. Waste time put the money in the bank.
went go and	He went go work. It went peep.
went	Sure, you went cry. You went go tell.
wild	My mama wild you know.

REFERENCES

1. ADAMS, R. Interracial Marriage in Hawaii. New York: Macmillan, 1937. Pp. 353.
2. ———. The Peoples of Hawaii. Honolulu: American Council Institute of Pacific Relations, 1933. Pp. 58.
3. ALEXANDER, W. D. A Short Synopsis of the Most Essential Points in Hawaiian Grammar. Honolulu: Thos. G. Thrum, 1908. Pp. 59.
4. BLAKE. A Grammar of the Tagalog Language. New Haven: American Oriental Society, 1923. Pp. xxxi+324.
5. BOYD, WM. The development of a child's vocabulary. *Ped. Sem.*, 1914, 21, 95-124.
6. BRANDENBURG, G. C., & BRANDENBURG, J. Language development during the fourth year. *Ped. Sem.*, 1916, 23, 14-29.
7. CARR, D. Polysyllabicity in Chinese. Lwow, Poland: Rocznik Orientalistyczny, 1934.
8. CHUN, ELINOR YUK LIN. A study of sentences by the bilingual child of Chinese ancestry in Honolulu. Honolulu: Univ. Hawaii, 1935. (Unpublished master's thesis.)
9. DUNN, J. A Grammar of the Portuguese Language. Washington, D. C.: National Capital Press, 1928. Pp. 661.
10. ENCYCLOPEDIA BRITTANICA. London, 1929. (14th ed., Vol. 5, pp. 567-570.)
11. FISHER, MARY S. Language Patterns of Preschool Children. New York: Teach. Coll., Columbia Univ., 1934. Pp. ix+88.
12. GATES, A. I. A Reading Vocabulary for the Primary Grades. New York City, Teachers College, Columbia University, 1926. Pp. 23.
13. HOFFMAN, M. N. H. The Measurement of Bilingual Background. (T. C. contrib. No. 623). New York: Teach. Coll., Columbia Univ., 1934. Pp. v+75.
14. INTERNATIONAL KINDERGARTEN UNION. A Study of the Vocabulary of Children Before Entering First Grade. Baltimore International Kindergarten Union, 1928. Pp. 38.
15. ISAACS, SUSAN. Intellectual Growth in Young Children. With an appendix on children's "why" questions by Nathan Isaacs. New York: Harcourt, Brace, 1930. Pp. xi+370.
16. LEE, A. K. A study of the Hawaiian vocabulary of certain groups of preschool children in Hawaii. Honolulu: Univ. Hawaii, 1937. (Unpublished master's thesis.)
17. LIND, A. Occupational trends among immigrant groups in Hawaii. *Soc. Forces*, 1928, 2, 290-299.
18. MASUOKA, J. Race attitudes of the Japanese people in Hawaii. Honolulu: Univ. Hawaii, 1931. (Unpublished master's thesis.)
19. MCCARTHY, D. A. The Language Development of the Preschool Child. Minneapolis: Univ. Minnesota Press, 1930. Pp. xiii+169.
20. MCGOVERN, WM. M. Colloquial Japanese. New York: Dutton, 1934. Pp. viii+234.
21. MERIGGI, P. Sur la structure des langues "groupantes". *J. de Psychol.*, 1933, 30, 185-216.

22. Miyamoto, S. A study of the Japanese language ability of the second and third generation Japanese children in a Honolulu language school. Honolulu: Univ. Hawaii, 1937. (Unpublished master's thesis.)

23. Nice, M. M. The speech development of a child from eighteen months to six years. *Ped. Sem.*, 1917, **24**, 204-243.

24. Piaget, J. The Language and Thought of the Child. New York: Harcourt Brace, 1926. Pp. xxiii+246.

25. Porteus, S. D., & Babcock, M. E. Temperament and Race. Boston: Badger, 1926. Pp. xiv+364.

26. Reinecke, J. Language and dialect in Hawaii. Honolulu: Univ. Hawaii, 1935. (Unpublished master's thesis.)

27. Reinecke, J. E., & Tokimasu, A. The English dialect of Hawaii. *Amer. Speech*, 1934, **9**, 48-58, 122-131.

28. Smith, Madorah E. Grammatical errors in the speech of preschool children. *Child Devel.*, 1933, **4**, 183-190.

29. ———. The influence of age, sex and situation on the frequency, form and function of questions asked by preschool children. *Child Devel.*, 1933, **3**, 201-213.

30. ———. An investigation of the development of the sentence and the extent of vocabulary in young children. Univ. Iowa Stud.: Stud. Child Welfare, 1926, **3**, No. 5. Pp. 92.

31. ———. A study of some factors influencing the development of the sentence in preschool children. *J. Genet. Psychol.*, 1935, **46**, 182-212.

32. Terman, L. M. Genetic Studies of Genius, Vol. I. Stanford Univ., Calif.: Stanford Univ. Press, 1925.

33. Underwood, H. G. An Introduction to the Korean Spoken Language. (Sec. Ed.) New York: Macmillan; Yokohama, Shanghai: Kelly & Walsh, 1914. Pp. 475.

34. United States Bureau of the Census. Fifteenth census of the United States: 1930, Outlying Territories and Possession. Washington, U. S. Government Printing Office, 1932. Pp. 338.

35. Williams, S. W. The Middle Kingdom. New York: Scribner's, 1899. (Vol. I, Chap. 10.)

36. Yee, Alyce. The development of language among bilingual children of Chinese ancestry in the city of Honolulu as measured by their use of English words. Univ. Hawaii, 1935. (Unpublished master's thesis.)

37. Zyve, Claire T. Conversation among children. Teach. Coll. Rec., 1927, **29**, 46-61.

$7.00 per volume
Single numbers $3.00

QUARTERLY

August, 1939
Volume 21, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota
Charlotte Bühler
Universität, Wien
William H. Burnham
Clark University
Cyril Burt
University of London
Leonard Carmichael
Tufts College
Ed. Claparède
Université de Genève
Edmund S. Conklin
Indiana University
Sante De Sanctis
R. Università di Roma
Arnold Gesell
Yale University
William Healy
Judge Baker Foundation, Boston
Leta S. Hollingworth
Teachers College, Columbia University
Buford Johnson
The Johns Hopkins University
Harold E. Jones
University of California
Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal College
K. S. Lashley
Harvard University
A. R. Luria
Medico-biological Institute, Moscow
Toshio Nogami
Kyoto Imperial University
Henri Piéron
Université de Paris
George D. Stoddard
State University of Iowa
Calvin P. Stone
Stanford University
Lewis M. Terman
Stanford University
Godfrey Thomson
University of Edinburgh
E. L. Thorndike
Teachers College, Columbia University
C. J. Warden
Columbia University
John B. Watson
New York City
Helen Thompson Woolley
Teachers College, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

Psychological Foun

Provincetown, Massachusetts
Copyright, 1939, by The Journal Press

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the Act of March 3, 1879.

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879.

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out. Beginning with January, 1938, the following methods and terms will be in force:

(1). The size of each volume will be 400-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinco or half tone cuts.

(6). The publication charges will be as follows:

4 page signature	$12
8 " "	16
12 " "	24
16 " "	32
20 " "	40
24 " "	48
28 " "	56
32 " "	64
36 " "	72
40 " "	80

These charges are tentative and will be adjusted as justified by the subscription list.

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Beginning with January, 1938, eight hundred to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $3.00

QUARTERLY

August, 1939
Volume 21, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

AUGUST, 1939

(Manuscripts are printed in the order of final acceptance)

DOMINATION AND SOCIAL INTEGRATION IN THE BEHAVIOR OF KINDERGARTEN CHILDREN AND TEACHERS 287
By Harold H. Anderson

THE CAPACITY OF THE RHESUS AND CEBUS MONKEY AND THE GIBBON TO ACQUIRE DIFFERENTIAL RESPONSE TO COMPLEX VISUAL STIMULI 387
By William Egleston Galt

Copyright 1939, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1939, **21**, 287-385.

DOMINATION AND SOCIAL INTEGRATION IN THE BEHAVIOR OF KINDERGARTEN CHILDREN AND TEACHERS*

HAROLD H. ANDERSON

Department of Psychology, University of Illinois

Acknowledgments 289

I. Introduction 291

II. The measurement of teachers' behavior 297
  - A. Aims 297
  - B. Methods 297
  - C. Subjects 300
  - D. Formulation of the categories 301
  - E. Definitions of the categories 305
  - F. Consistency of observations by two observers . . . 309
  - G. Consistency of the time samplings 320
  - H. Differences between Teachers A and B 328
  - I. Differences in contacts during the morning and afternoon for the same teachers 330
  - J. Number of teachers' contacts per hour 331
  - K. Sex comparisons in teachers' contacts with individual children in School *X* 349
  - L. Correlation of teachers' dominative contacts with teachers' integrative contacts 350
  - M. Relation of teachers' contacts to mental age and chronological age of the child 351
  - N. Summary and conclusions 352

III. Domination and integration in the social behavior of kindergarten children in an experimental play situation 357
  - A. Aims 357
  - B. Methods 357
  - C. The portable testing booth 358
  - D. Subjects 358
  - E. Methods of scoring 359
  - F. Definitions 359
  - G. Consistency of observations by two observers . . . 363
  - H. Sex differences 363
  - I. Domination-integration ratios 373
  - J. Consistency of domination and integration scores between first and fifth pairings 374
  - K. Correlations of child's domination and integration scores with teacher's contacts and other factors 375
  - L. The dynamic nature of interactivity as shown in domination and integration scores 377
  - M. Summary and conclusions 378

References 385

---

*Received in the Editorial Office on April 3, 1939, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

OWLEDGMENTS

The studies reported herewith were undertaken with the assistance of the George Davis Bivin Foundation and the Research Board of the University of Illinois. Acknowledgment is gratefully made to Professor Herbert Woodrow, Head of the Department of Psychology, for valued suggestions and encouragement and to Joseph E. Brewer and Dorothy Walker Loeb, research assistants. Gladys Lowe Anderson, Helen M. Brewer, and Mary Frances Reed assisted at several stages of the study in giving mental tests, making observations, or collecting data. Preliminary observations were made in a number of schools and institutions in Iowa and Illinois. The final data reported here, however, were obtained through the coöperation of Superintendent T. H. Cobb, the principals, and teachers of the Public Schools, Urbana, Illinois, and of Superintendent V. L. Nickel, the principals, and teachers of the Public Schools, Champaign, Illinois. The writer is indebted also to E. L. Welker for statistical assistance.

HAROLD H. ANDERSON

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

I. INTRODUCTION

The present report on domination and integration in the environment of kindergarten children is an outgrowth and extension of previous research. This report presents two research projects that may be treated as separate investigations. They are both founded, however, on the same psychological concepts and are constructed on the same basic hypotheses regarding human behavior.

Theoretical considerations underlying this research have been presented elsewhere (2, 3, 4, 5, 6, 7, 8). Briefly, integrative behavior is a growth phenomenon that involves a change in structure or function of an organism through the confronting of differences. Integrative behavior may be said to be a voluntary or spontaneous abandoning of psychological structure (concepts) or function (goals, purposes, activities) which is another way of describing the changes commonly called learning. Integrative behavior is characterized by a voluntary or spontaneous *yielding* or abandoning of the existing structure or function for a new structure or function that is in the process of becoming. Yielding is a spontaneous response to difference. It is the response of a secure person, of one who is free to abandon himself to become something new and different. Growth is commonly regarded as the amount or degree or nature of change in structure or function. Integration is the technique of responding to differences by which growth is achieved. The degree or amount or extent of integration occurring during a given time interval in the interplay between an individual and his environment may be regarded as a measure of the rate of growth, the rate of change in structure or function.

Growth is to be observed and measured along a theoretical continuum; it is not an all or none phenomenon; it is relative and operational. Growth is process; it is dynamic, not static. The cross-sections of this process which an experimenter observes and records or measures are at the moment of recording already historical components of the individual's past experience. Life moves on as a process of responding between the individual and his environment. Response is by definition change: the individual is different for having responded. Growth is creative (differentiative); it is characterized by the emergence of originals, originals in structure or function.

It is not possible to define optimum growth nor is it necessary in order to study growth. What one always studies under physical, mental, or personality growth is the something-less-than growth at its best, the responding at a point somewhat short of the upper theoretical terminus of the continuum between atrophy, stagnation, or disintegration at one end and optimum growth at the other.

Growth or learning occurs only through the confronting of differences; growth and learning are phenomena of yielding to differences in which the organism emerges with a changing structure or function.

Integrative behavior. Personality growth or integrative behavior is a process of change in structure or function that results from increasingly complex relations with persons different from one's self. Personality growth occurs when the individual is able to seek and to discover or define common purposes in activities that are mutually satisfying to himself and to those in his environment. Personality growth or development is directly proportional to the spontaneity (security or absence of fear) with which a person responds to his environment.

Integrative behavior is flexible, dynamic, yielding, spontaneous; it shows no fear of abandoning status, no fear of change. An integrating person seeks and finds common purposes with another; he expends energy with another, not against another.

Integrative behavior in one person induces integrative behavior in another. An integrating person accepts another *as he is*, thus contributing to the other's security, and thus making it possible for the other person to be spontaneous, to be himself. There is no greater psychological security than for one to be accepted as he is. In a child guidance clinic the integrative relationship is called rapport.[1]

---

[1] Compare a statement by a psychiatrist describing a relationship which he calls "specific and unique." Frederick H. Allen, writing on *The Child as the Therapist Sees Him* says: "First, I might indicate the nature of the therapeutic relation with a child. . . .

"This is a specific and unique relation established with a child who is presenting some difficulties in his behavior. Usually he is brought and has had little to say as to whether this is a thing he wants to do. From the beginning the therapist tries to give the child the feeling that here he is taken just as he is. He is accepted as a child, not as a problem, with no attempt to mould him into anything other than what he can be at the moment. . . .

"The therapist is an unusual adult in the child's experience because the child finds himself accepted by him exactly as he is. There is no other person present. . . . There is no attempt to make him into something else, *Dominative behavior.* Dominative behavior on the contrary is rigid, fixed, static. A dominating person has his mind made up, has his goals or desires predetermined. He does not yield to differences; he is not abandoning his status; he is trying to preserve status. He is not seeking a better understanding of another nor is he trying to achieve a redefining of desires, values, or objectives in order to discover a lower common denominator of differences. He is expending energy against another. He is not reducing conflict, he is either maintaining or increasing the conflict of differences.

Domination seeks first not self-abandoning, but self-preservation. It resists change. However justified the individual may be in preserving the status quo, domination is decidedly less than growth at its optimum.

Resistance and submission. Domination tends to induce resistance. But resistance is itself dominative behavior. If the relative strength is too great, domination will produce submission.

Resistance and submission are both fear responses; they show fear of losing status quo. Submission is not to be confused with yielding. Submission is a coerced response of fear; it is not spontaneous. Yielding is a spontaneous response by an individual so unconcerned with his own security that he is not afraid to change or give up his present status. Resistance and submission are responses to domination; they are both something-less-than spontaneous, and both something-less-than expression of growth.

Domination and integration are different and unrelated techniques of responding to differences. Domination can obstruct integrative behavior where the intolerance of difference and the balance of power are both too great. Domination, being the negation of growth, cannot induce integrative behavior. Integrative behavior creates no situation in which another person feels the need to maintain or protect his status; integration thus does not induce domination.

There is probably no situation in which the responding process

---

nor to change his pattern of behavior. The child is neither scolded nor blamed, nor is he told what to do. . . ." (1, p. 3-4).

In other words the therapeutic relation has an absence of domination and an abundance of integration. The psychiatrist is seeking a common purpose with the child, is expending energy with the child, not against him. Integrative behavior of the psychiatrist (accepting the child as he is, yielding to differences) induces integrative behavior in the child. When the induction works, psychiatrists call it rapport.

> between an individual and his environment is characterized uniquely by domination or by integration; the phenomena of relating are too complex. There are many momentary situations, however, in which one of these types of responding can be said to be preponderant, in which energy is spent with another or is spent against him, in which efforts are made to maintain or increase conflict of differences or to reduce conflict of differences.
>
> The present study does not include data on submission. It does include data on domination and resistance to domination, and on integrative behavior or the expenditure of energy in the direction of common purposes (6, pp. 344-346).

In the study of the social interplay of children of preschool age in an experimental play situation from which the above quotation was taken, evidence was offered in support of the following hypotheses:[2]

> 1. Domination incites domination (resistance).
> 2. Domination is dynamically related to dominative behavior in a companion.
> 3. An insecure child makes a companion insecure.
> 4. Energy expended against a companion will induce energy in the companion directed against oneself.
> 5. Domination does not induce integrative behavior in a companion.
> 6. Domination is not only different from, but where a potential avenue of escape is left open it is dynamically unrelated to integrative behavior.
> 7. An insecure child does not make a companion secure.
> 8. Energy expended against a companion does not induce energy in the companion directed in a common purpose with oneself.
> 9. Integrative behavior in a child induces integrative behavior in the companion.
> 10. Integrative behavior in a child is dynamically related to integrative behavior in a companion.
> 11. A secure child makes for security in a companion.
> 12. Energy expended in a common purpose with a companion induces energy in the companion directed with one's own purposes.

---

[2] The following hypotheses together with the supporting evidence are presented in a summary form in Table 26 (6, pp. 400-401).

13. Integrative behavior does not induce dominative behavior in a companion.

14. Security in a child does not induce insecurity in a companion.

15. Energy spent in a common purpose with a companion does not induce the expenditure of energy in the companion directed in opposition to oneself.

Both logic and experience tend to support the application of the above hypotheses to the relations between teachers and children. Since the relations between teachers and children are psychologically more complex and involved than the child-child relations at the preschool level, the development of measures of domination and integrative behavior in teachers' responses to children appeared to be a major problem in itself. It seemed, however, that the problem of measuring teachers' contacts with children would offer fewer difficulties at the younger age levels of school children than at the older ages. Also it has seemed important to determine whether child-child relations at the school ages would confirm the findings with children of preschool age.

Accordingly two investigations were undertaken at the kindergarten level. Part II will present the report on the measurement of domination and integration in the behavior of kindergarten teachers. Part III will present the report on the measurement of domination and integrative behavior in the social interplay of kindergarten children in the same kind of experimental play situation used in the previous study of children of preschool age.

II. THE MEASUREMENT OF TEACHERS' BEHAVIOR

A. AIMS

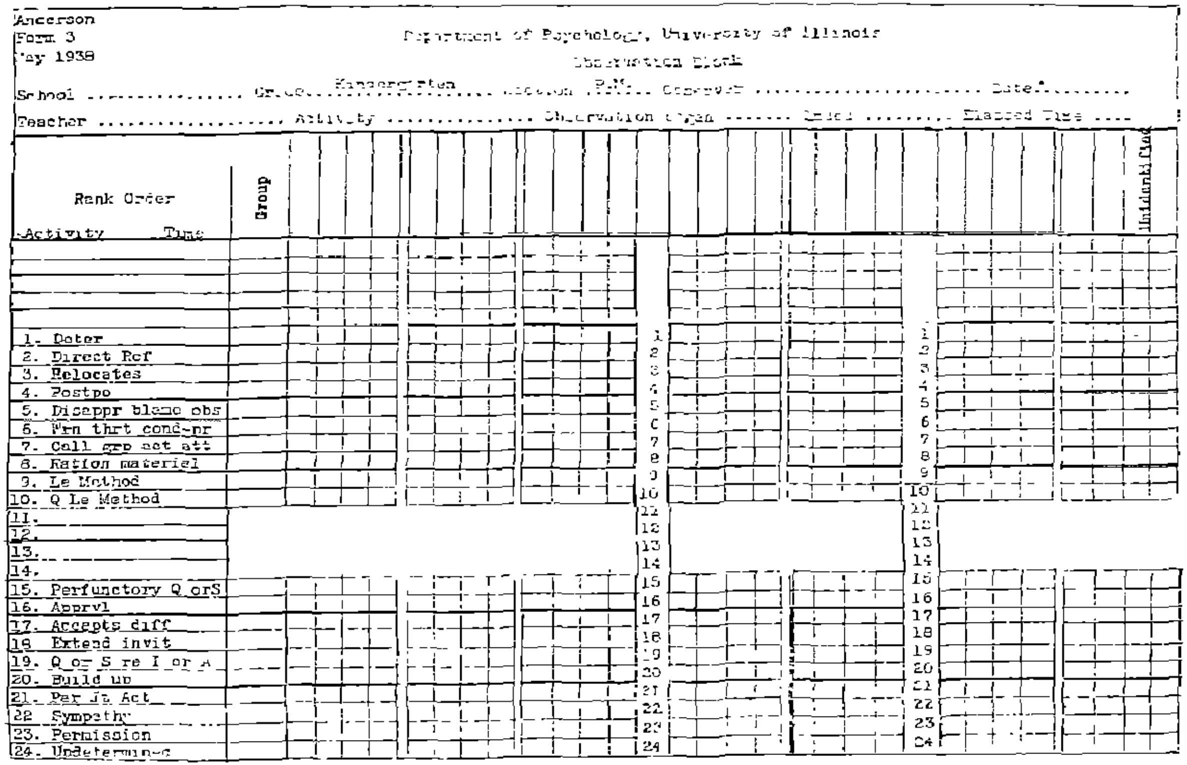
The purpose of this part of the study was to develop in terms of the concepts of domination and integrative behavior quantitative measures for observable contacts which teachers had with kindergarten children.

B. METHODS

It was believed in the early stages of the study that for reasons of practical economy in collecting data an experimental situation would have to be devised in order that the interplay between the teacher and the child be speeded up. One such situation seemed quite hopeful: a great deal of teacher-child interplay was noted during a half hour in which the children gave a play which they had been instructed to plan and to produce without adult assistance. This situation offered experimental possibilities for observing and measuring the degree to which children were allowed to be "spontaneous," allowed to "be themselves," and the extent to which with a defined minimum of adult expectations, the demands of the teacher made incursions on the imaginative work of the children. Such situations could be set up experimentally at all age levels. It was found, however, that the frequency of teacher-child contacts in the actual kindergarten situation was so great as to make an experimental situation unnecessary. The method adopted was that of direct observation in the kindergarten schoolroom.

From files of running notes taken during the past five or six years describing teacher-child interplay and from new observations made specifically for this purpose, a set of preliminary categories of teachers' contacts was constructed. Whenever a contact with a child failed to fit any of the provisional categories, a new category was defined. Thus over a period of several weeks of meticulous examination of notes and discussion of the categories a preliminary observation blank and set of definitions of contacts was devised. Then the blank was tried out by two observers recording simultaneously in the schoolroom, following which the definitions of categories were further discussed and the observation blank revised.

The observation blank was devised to contain five minutes of observations. Each blank bore the identifying information showing



Anderson
Form 3
May 1938

Department of Psychology, University of Illinois
Observation Blank

School .... Grade ..Kindergarten...... Session ..P.M... Observer .... Date.........
Teacher .... Activity .... Observation began .... Ended .... Elapsed Time ....

Rank Order
Activity Time
Group
Unidentified

1. Deter
2. Direct Ref
3. Relocates
4. Postpo
5. Disappr blame obs
6. Wrn thrt cond-pr
7. Call grp act att
8. Ration material
9. Le Method
10. Q Le Method
11.
12.
13.
14.
15. Perfunctory Q orS
16. Apprvl
17. Accepts diff
18 Extend invit
19. Q or S re I or A
20. Build up
21. Par in Act
22 Sympathy
23. Permission
24. Undetermined

FIGURE 1

the school, grade, section, date, observer and teacher and in addition the name of the activity in which the group was engaged, the time the observation period began and ended, and the elapsed time of the observation period. The blank which was adopted after experimentation with two other forms is shown in Figure 1. It had the names of the children at the tops of vertical columns and the names of the categories of teacher contacts on the horizontal rows. It will be noted that there are no categories for Nos. 11 to 14 inclusive. These numbers designated categories on previous experimental forms of the observation blank which were finally combined with other categories. Because the experimenters had memorized the other categories by number and by relative position on the blank, the numbers and original spacing were retained on the final observation blank. For aid in recording categories quickly two additional columns of guide numbers were inserted in the blank.

1. *The Schoolroom Situation*

The kindergarten groups in School *X* used a large room divided by low cupboards built perpendicular to the side walls. There was thus a large play room at one end and a smaller room at the other. The smaller room contained tables and was used for hand work and lunches. The head teacher, designated Teacher *A*, took the leading rôle with the children much more frequently than did the assistant teacher, who is designated Teacher *B*. The assistant, however, had charge of the music period, at which time the head teacher usually withdrew to arrange hand work for the group. Most of the contacts which Teacher *B* had with the children occurred during the music period. The kindergarten in School *Y* was conducted in a single large room. All observations of teachers' contacts were taken in these rooms.

2. *Methods of Observation*

The observers were instructed to observe the teacher who was playing the major rôle with the children. The frequencies of teacher contacts thus represent those of the one teacher most active at the time, but by no means all the contacts which the children received from both teachers during the observation period.

An attempt was made to limit the observations on one blank to a period of five minutes. When in the midst of an arbitrary five-minute period the leading teacher turned the group over to the other teacher, or when there was an abrupt change in the activity of the group, as from music to handwork, the observer began recording at once on a new blank.

Moreover, because of the rapidity of interplay between the teacher and the children, observers occasionally ran inadvertently over the five-minute period before taking a new observation blank. Because of the rapidity of interplay and the consequent necessity for constant attention, observers did not observe continuously for more than an hour of consecutive observation periods.

Observers sat on low kindergarten chairs at the side of the room where they were relatively inconspicuous and at the same time near enough to hear and to see. The observation blanks were fastened to a clipboard and held on the lap.

The observer marked the blank by placing in the child's column one tally for each contact which the teacher had with that child individually. If the contact was directed to the group rather than to an individual the tally was recorded in the *"Group"* column.

Where the teacher chose several children successively for a common activity the *"Rank order"* was recorded, the observer writing the name of the activity (e.g., singing, rationing materials, etc.) and the time, and marking the blank not by tallies, but by serial numbers under the respective children's columns.

If the teacher made some contact with a child or with the group, but the nature of the contact was not clear the observer recorded a tally under Category 24, *"Undetermined."* If the nature of the category was clear, but it was not known with whom the contact was made, the tally was placed in the *"Unidentified"* column at the right-hand side of the blank. This column collected not only a few contacts which occurred when the observer for one reason or another failed to see or to hear, but a number of such partially disguised contacts as *"Some little boy forgot to remember what we said about hands—or eyes—or feet."*

C. SUBJECTS

The subjects were 55 children attending three different kindergarten groups in two school buildings and taught by three teachers. Teachers *A* and *B* taught both morning and afternoon groups in School *X*; Teacher *C* taught a morning group in School *Y*.

The kindergartens, although housed in public school buildings, were privately supported. There was a tuition fee of $3.50 per month per child. The children came from a university community from families able and willing to pay the fee. In general the children were superior in intelligence. In School *X* an attempt was made to enroll younger children in the morning group.

Table 1 gives the numbers of children by sex and groups, together with means and ranges of chronological age, mental age, and *IQ*. It will be noted that in all groups the girls only slightly outnumbered the boys, there being a total of seven more girls than boys. The 11 children in School *Y* had a mean chronological age six months below that of the younger morning group of School *X*, and 16 months below that of the afternoon group of School *X*. The morning group of School *X* contained one girl who was the oldest and who had the lowest *IQ* in the entire enrollment. For this reason separate figures have been given in Table 1 including and excluding this child.

TABLE 1

NUMBERS OF CHILDREN BY SEX AND BY SCHOOL, WITH MEANS AND RANGES OF CHRONOLOGICAL AGE, MENTAL AGE, AND *IQ*

School	Session	Boys	Girls	Total
X	A.M.	10	13	23
	P.M.	9	12	21
Y	A.M.	5	6	11

		Chron. age		Mental age		*IQ*	
School	Session	Mean	Range	Mean	Range	Mean	Range
X	A.M.	62.7	54-81	72.3	53-93	117.2	70-156
	*A.M.	61.9	54-73	73.1	53-93	119.1	75-156
	P.M.	71.9	66-76	87.8	70-115	122.6	96-156
Y	A.M.	55.7	40-66				

*One child in the morning session had the highest *CA* of the group (81 months) and the lowest *IQ* (70). The next higher *IQ*'s were respectively 75 and 99. The first set of figures given for the morning session include this child. The second row of figures for the morning session show means and ranges without the distortion which this one child introduces.

D. FORMULATION OF THE CATEGORIES

The designations of categories on the observation blank are abbreviated to the point of being meaningless without the definitions. The full titles of the categories together with examples of teachers' contacts which they include are given below. Actual observations of teachers' behavior constituted the basis on which each category was *constructed and defined.*

It scarcely need be pointed out that these definitions and categories were arrived at with great difficulty. In the previous study of dominative and integrative behavior (6, pp. 355-358) it was found that teachers were better able to rate children for their dominative behavior than for their integrative behavior. The question was raised whether the nursery school teachers in that study were not still "problem conscious."

In the present study the preliminary categories of teachers' contacts included, as might have been expected, a number of examples of extreme forms of domination and of integrative behavior. When it came to applying these categories to the kindergarten situation in School *X*, the observers were plainly baffled. The teachers were soft-spoken, attentive, patient, considerate of the children. There was complete absence of anything impetuous or impatient in the teachers' responses to the children. The observers were at a loss as to how to record anything that was dominative or "unhygienic." Teachers, to be sure, told children to do certain things and not to do other things. But all teachers do that, and not to do so in a casual, sympathetic way seemed too much to expect. The experimenters became aware that they were constantly evaluating the teachers' responses to the children, and practically everything they saw seemed acceptable.

If a teacher in introducing the music period said, *"I'm going to sing you a song,"* the observers felt that that was definitely a social contact with the group, but it was not clear whether it could be checked as dominative or integrative or what difference it made what one called it. Only by gradual stages in the preliminary steps of the investigation were the observers able to see the contacts as descriptive units and to abandon for the purposes of observation the interpretive or evaluative point of view. Only after much labor were the experimenters able to devise criteria and arrive at definitions that would record domination even if expressed in a "soft voice." If measurement were to be preceded first by careful observation it did seem that a key to the evaluative difficulty could be found by checking the teachers' remarks against the criteria of conformity versus participation. Did the teacher tell them or ask them?

Did she base decisions on her own desires or judgment or did she allow some measure of interplay for the child's desires, the child's judgment?

It did seem as though for each isolated remark it made little difference whether the teacher told them that she was going to sing a song or asked them whether they would like a song or, if so, what song would they like to hear? It seemed logical to expect, however, that an accumulation of tallies that would record such simple differences would make a distinction between some teachers and others, that in some schoolrooms there would be a great deal or a preponderance of teachers' contacts in which the teacher *told* the children what to do, what she was going to do, or what the activities were to be. And on the other hand, it was conceivable that other teachers would have much lower frequencies of such techniques and might perhaps be found to be giving the pupils a proportionately higher number of opportunities to use their own judgments. No presumption was made as to whether these two techniques might be inversely related. It was found among preschool children (6, pp. 397-398) that in the experimental situation used, domination had a zero relation with integrative scores. With high domination scores in one child one could predict high resistance (domination) or counter-attack in the companion, but he could not predict the companion's integrative score nor the child's own integrative score. There was not an inverse, but a zero correlation between domination and integration.

Domination in the present study includes social contacts in which the activity of the child or of the group is determined out of the experience or judgment of the teacher. Such a contact is psychologically different from the contact in which there is a democratic interplay, in which the determination of the child's activity comes from a broader experiential base that includes the judgment or choices of the child himself. The psychological assumptions are that the child "learns" less arithmetic if father does all his problems for him, and he grows less in other respects to the extent that the teacher decides what is to be done and how and when to do it. Telling them is assumed to be not only psychologically different from asking them, but in general it is assumed to be less propitious for growth, learning, and problem-solving.

Categories of teachers' contacts 1 to 8 inclusive are classified as domination; Categories 15 to 23 are classified as integrative behavior; Categories 9 and 10 which had low frequencies were regarded as ambiguous hybrids, not clearly classifiable as domination or integration. It is believed, however, that the majority of contacts checked in Category 9 belongs more properly in the group of dominative techniques, and that contacts checked in Category 10 would fall more properly among the integrative contacts.

There was evidence of some difficulty in checking items in separate categories within the respective groups of dominative and integrative techniques. There was less tendency for one observer to record a tally as domination and for the other observer to record the contact as integrative behavior. In other words one observer might check a contact in Category 5 (e.g., disapproval or blame) while the other might have checked it as Category 6 (warning or threat), but there seemed to be less probability that a contact checked by one observer in one of the categories 1 to 8 would be checked by the other observer in one of the categories 15 to 23.

The purpose of this study was to develop measures in frequencies for the dominative and integrative contacts which teachers had with children. It is not yet time to attempt to evaluate these contacts or to compare one contact with others in its dominative or integrative grouping. It seems obvious that some contacts are more dominating than others and that some are more integrative than others. No attempt is made here, however, to measure degrees or intensities or relative strengths of these contacts, nor the relation between these contacts and the child's growth or mental hygiene. It is believed, however, that these problems are amenable to measurement and to scientific determination. The establishing of categories was an arbitrary matter of convenience in recording the teachers' contacts, and also a means for a preliminary search for more refined analyses of teachers' behavior. Analyses of the data have been made to show the consistency with which two independent observers recording simultaneously were able to record teachers' contacts by categories. But in the treatment of the data to show the contacts of the teachers with the individual children and with the group all the dominative and all the integrative categories respectively are combined.

E. DEFINITIONS OF THE CATEGORIES

Rank order: Choosing child to sing alone or in pairs; choosing child to line up in specified order; detention,—"all may go except—."

1. *Determines a Detail of Activity or Acts for the Child in Carrying Out a Detail*

Includes instances where *T* (teacher), in order to rush through to an end, goes ahead and does things for the child.

T: "I'm going to sing you a song."
"You listen so you will know it."
"You will have to fold yours like this."
"I'll get you another handle."
"I'll put the names on the baskets."
"All right, we won't play that game any more."
"I thought you would like to make some baskets to take home" (one tally).

2. *Direct Refusal*

T answers *"No"* to a direct request.

3. *Relocating, Reseating, or Placing Children in Different Relation to Each Other or to Property, i.e., Different from the Relation which the Children Have Themselves Selected*

T: "Henry, Janet, Sam, please sit down."
(Later—*T:* "I don't want to speak to Henry, Sam, and Janet again." Check each child for No. 6, Warning, threat.)

4. *Postponing, Slowing Up the Child*

T: "Not now."
"Wait just a minute."
"Later on."
"Sometime soon."
"Tomorrow."
"In a few minutes."

Holds back the fast ones.

Obstructs differentiation, originality, individual differences, variability within a group.

T: "Betty Lou, go back and wait until I come around."
"Wait at your place until I give you one."

(*Note:* The question as to how much waiting or "organization" was necessary for the conduct of the group gave the experimenters some difficulty at first. But the question was an ethical question, not scientifically descriptive. The objective problem was what could be observed and recorded reliably. In some groups there is more holding back of the fast ones than in other groups. That much can be descriptively determined. Then the question: How much is a "good thing" or how much is "justified" or even "necessary" can depend on defined pedagogical objectives or ends or purposes for having schools.)

E.g., Sally had just finished pasting. *S: "I've got to wash my hands."* The toilet and wash room were just behind the place where she was sitting. It was built for the special use of the children in that particular room and was *inaccessible* from outside the room.

T: "Just a minute, Sally, and we will all go and wash our hands." Following which, S was obliged to stand in line and *wait her turn.*

5. *Disapproval, Blame, or Obstruction*

"Hurry up" implies disapproval.

"I'm waiting."

"One little boy—I don't see his eyes at all."

Check "unidentified."

E.g., Pete was not listening, did not know what *T* had said;

T: "I didn't think you would know."

E.g., *T* had asked them all to place their finished baskets in the middle of the table and not to play with them.

T: "Somebody forgot not to play with their basket." (Check the unidentified.)

(*Note the overlapping between the holding back of the fast* ones in No. 4 and the stronger environmental pressure of obstructing with an added implication of blame or disapproval in No. 5.)

6. *Warning, Threats, or Conditional Promises*

T: "I don't want to speak to Henry, Sam, and Janet again."

T: "Now if we all sit nicely and keep our hands to ourselves, we might have two stories."

7. *Call to Attention or to Group Activity*

E.g., the bugle call: sounding a certain chord on the piano.

T: "Girls and boys ———."

T: "Let's see who is listening."

8. *Rations Material*

T makes decisions as to amount, kind, etc., e.g., amount of paste, amount of grass for rabbit nests.

(Implication is that rationing of materials is psychologically more than an administrative convenience; it deprives the child of an opportunity to exercise his own judgment, to decide for himself how much it will take for the job at hand; and for this reason it is an expression of *T* domination.)

9. *Lecture Method*

T gratuitously defines a problem or anticipates the question and gives the answer. (The "sez you" category.)

E.g., *T*, passing out paper;

T: "The paper is to keep the paste off the tables."

(If there was a problem of keeping the paste off the tables, the children might have contributed from their experience *in defining the problem, especially since only the children* got paste on the table. As a matter of fact, the tables were made so paste could be washed off. Paste actually got on the tables, and as a later part of the routine a child with great enthusiasm did wash the tables after the children were through pasting.)

T: "You won't need your scissors" (check No. 9).

(But) "Don't get your scissors" (check No. 1).

10. *Questions: Lecture Method*

Questions where the answers are in the back of the book or in the teacher's experience.

T: "What did the birdie say?"

If there is only one answer, then check No. 10. If the child is permitted to give an imaginative answer, then check under No. 19 or No. 20.

(Eleven to fourteen inclusive deleted on the blank.)

15. *Perfunctory Question or Statement*

Indifferent "Thank you's":

T: "Isn't that interesting?"—a bare response, but a response never the less.

T: "Is that so?"

Disregard perfunctory remarks when they are combined with something else representing a stronger response, e.g., child has finished basket. Says: "Isn't my basket cute?"

T: "Yes, but we haven't got it all pasted. The decorations are not on." The *"Yes"* is regarded as perfunctory, but not checked. The remainder of the statement is checked as No. 5, Disapproval.

16. *Approval*

Includes rewards, prizes, competitive favors.

T: "I think that's fine."
"That's fine."
"Billy's row is standing the straightest."

17. *Accepts Difference*

Observer must be alert for negative votes, declinings, expressions of difference, conflicts of difference. Whenever the *T* makes an offer or gives an invitation, and the child declines, some category should be checked for the *T*'s response: She either accepts the difference (No. 17); or she reproves (No. 5); or she renews her request (No. 18).

E.g., *T:* "Jimmy, would you like to sing this one (song) up here? (beside *T*)" Jimmy declines.

T turns to another child.

(Check rank order for Jimmy, No. 18, Extends invitation; check the other child, rank order for No. 18; check Jimmy for No. 17, Accepts difference.)

18. *Extends Invitation to Activity*

"Who wants to be a pony?"
"Who would like to be a robin?"

Call for a show of hands. The choice rests with the children. It must be obvious that there is no element of exhortation and that a child can still decline. Under a few circumstances will an *invitation* be made more than twice without obvious attempts to exhort; in which case check No. 1. A teacher's contact in category No. 1 cannot be declined without further exhortation or disapproval.

19. *Question or Statement Regarding Child's Expressed Interest or Activity*

Carries no presumption of opposition, antagonism, disapproval or urging.

"Dickie, are you waiting for paste?"
"How are you getting along?"

Includes the ice-breaker conversation.

"Do you have a dog at home?"

20. *The Build-up*

Highly integrative behavior.

Includes instances where *T* helps child to arrive at a better definition of a problem or a better solution, without giving the final answer.

If the final answer is given, the teacher's response is checked *as No. 9, or No. 10.*

T: "Look at your feet. Are you sure you have the right foot? Is your rubber on the right foot?"

21. *Participates in Joint Activity with Children*

Offers help, offers to participate.

Children playing ball. Ball rolls over near *T* who returns it.

22. *Sympathy*

T: "I'm sorry you hurt your finger."

23. *Permission*

T grants permission to child's request.

E.g., "May I get a drink?"
"May I pass the cookies?"

F. CONSISTENCY OF OBSERVATIONS BY TWO OBSERVERS

Because this study attempted to measure certain phenomena of human interplay that have hitherto been included among the "intangibles" of human relationships the problem of consistency of observers became at this stage an end in itself. Without sufficient consistency the desired measures would be valueless. To what extent and in what respects can two independent observers show consistency in the use of these measures when observing simultaneously?

Table 2 gives the time of observations made simultaneously by two observers during 73 observation periods in School *X* totaling 342.5 minutes, and 13 observation periods in School *Y* totaling 65 minutes. Table 2 does not include previous observations during which consistency of observers was not satisfactorily established, nor subsequent observations made for collecting the data that constituted the main body of the study.

1. *Consistency in Observing a Contact and in Speed of Recording*

It was noticed during the course of the investigation that the teachers' contacts with children occurred at times at such a rapid

TABLE 2

TIME OF OBSERVATIONS OF TEACHER CONTACTS MADE FOR DETERMINING DEGREE OF CONSISTENCY OF TWO OBSERVERS

Date	Teacher	Session	Obs. time (min.)	Between
		School X		
5/16	*A*	A.M.	47.0	9:51—10:49½
5/17	*A*	A.M.	28.5	9:30—10:22½
5/20	*A*	A.M.	24.0	10:24—10:50
5/16	*A*	P.M.	12.0	1:57— 2:37½
5/17	*A*	P.M.	18.0	1:37— 1:56½
5/20	*A*	P.M.	11.0	1:59— 2:17
5/23	*A*	P.M.	64.0	1:27— 2:57
5/16	*B*	A.M.	3.0	9:53— 9:59
5/17	*B*	A.M.	26.0	9:20—10:10
5/20	*B*	A.M.	27.0	9:59—10:52
5/16	*B*	P.M.	24.0	2:02— 2:35½
5/17	*B*	P.M.	15.5	2:00— 2:15½
5/20	*B*	P.M.	23.5	1:35— 1:58½
5/23	*B*	P.M.	19.0	1:51— 2:10
		School Y		
5/23	*C*	A.M.	65.0	10:20—11:25

rate as to raise a serious doubt whether the observers were able to record the contacts as fast as they occurred. If such were the case, then individual differences in the observers in defining the contacts would almost certainly appear. There would be fluctuations in the numbers of tallies on pairs of observation blanks or there would be considerable discrepancy in the total numbers of teachers' contacts recorded over an extended series of observation periods.

Seventy-three pairs of consecutive and simultaneous records of Observers *M* and *N* were analyzed. All the tallies on each observation blank were totaled and these totals correlated for the two observers. Table 3 gives the coefficients of correlation for separate

TABLE 3

COEFFICIENTS OF CORRELATION BETWEEN OBSERVERS *M* AND *N* FOR TOTAL NUMBERS OF CONTACTS PER OBSERVATION PERIOD FOR SEVENTY-THREE OBSERVATION PERIODS

Teacher	*A* & *B*	*A* & *B*	*A*	*B*	*A* & *B*
Session	A.M.	P.M.	A.M. and P.M.	A.M. and P.M.	A.M. and P.M.
r	.95	.96	.97	.96	.96
PE_r	.01	.01	.01	.01	.01
N	35	38	44	29	73

combinations of observation periods showing respectively contacts of Teachers *A* and *B* combined for the morning and again for the afternoon, contacts of Teacher *A* for the morning and afternoon combined and likewise for Teacher *B*, and contacts for all of the 73 periods combined. The lowest number of observation periods was 29 for Teacher *B* for both morning and afternoon combined.

All the coefficients are not only high but practically identical. They indicate that as far as the speed of recording was in question there was virtual identity in number of tallies recorded per observation period. This consistency of speed is shown whether the observations are for one teacher or the other; or whether they are made during the morning activities or during the afternoon program. These coefficients show also a very high agreement between observers in using the definitions of a "teacher contact."

The coefficients in Table 3 show that from one observation period to another there was little fluctuation in relative speed of recording. They do not in themselves show that the observers were recording at approximately the same speed. One observer could, for example, record at a rate twice as fast as the other, and if there were no fluctuations in relative speeds a perfect correlation would be possible. The check against the relative speed of the observers is found in the total number of tallies for the 73 observation periods. *During these periods covering five hours and forty-two minutes of* simultaneous observation, Observer *M* recorded 1,897 teacher contacts and Observer *N* recorded 1,893.

2. *Consistency in Recording Contacts with Individual Children*

A question next asked was: could the observers show consistency in observing with which child individual contacts were made? Could the observers show consistency in recording for each child the total number of contacts defined as domination, the total number of contacts defined as socially integrative, and the total number of all kinds of teacher contacts? To show consistency in these respects would mean that although the observers might confuse the categories within the respective groups of domination and integrative contacts or within the total of teachers' contacts, they would not show confusion in recording for individual children or in distinguishing between dominative and integrative contacts.

On each observation blank each child's column was totaled to

TABLE 4

Coefficients of Consistency of Observers for Teacher Contacts per Child with Categories Grouped
(One tally on the scatter diagram represents, e.g., tallies in Categories 1-8 for one child.)

School	Teachers	Session	Groups of categories									Total time (min.)	Mean time of correlated periods (min.)
			Total contacts 1-24			Domination 1-8			Integration 15-23				
			r	PE_r	N	r	PE_r	N	r	PE_r	N		
X	*A* & *B*	A.M. & P.M.	.94	.00	688	.93	.00	574	.46	.03	271	342.5	4.69
	A & *B*	A.M.	.94	.01	319	.94	.01	259	.44	.05	111	155.5	4.44
	A & *B*	P.M.	.94	.00	369	.93	.01	315	.46	.04	160	187.0	4.92
	A	A.M. & P.M.	.89	.01	411	.84	.01	319	.48	.05	191	204.5	4.65
	B	A.M. & P.M.	.96	.00	277	.95	.00	255	.50	.06	80	138.0	4.76
	A	A.M.	.87	.01	197	.87	.01	148	.53	.05	91	99.5	4.52
	B	A.M.	.97	.00	122	.97	.00	111	.10	.23	20	56.0	4.30
	A	P.M.	.91	.01	214	.80	.02	171	.44	.06	100	105.0	4.77
	B	P.M.	.96	.01	155	.94	.01	144	.53	.06	60	82.0	5.13
Y	*C*	A.M.	.89	.01	131	.87	.02	114	.29	.08	66	65.0	5.00

show the sum of tallies in Categories 1-24 inclusive, and subtotals were obtained for Domination Categories 1-8 inclusive and *for Integrative Categories 15-23 inclusive. In these computations* the two columns indicating *"Group"* contacts and *"Unidentified"* contacts were each treated statistically the same as a column for one child. Coefficients of correlation were computed for corresponding pairs of totals from the records of the two observers. These coefficients are shown in Table 4, together with the probable errors of the coefficients and the respective numbers of tallies from which the coefficients have been computed. This table gives also the total time in minutes and the mean time of the individual observation periods for each coefficient of correlation.

A glance at the first column of Table 4 shows that the coefficients range from .87 to .97. This column represents the total of all teacher contacts with each individual child recorded during one observation period by Observer *M* correlated with similar data as recorded by Observer *N*. For this kind of data these coefficients are very high. It may be noted that in all three cases where Teachers *A* and *B* are considered separately the observers show a higher agreement in recording the contacts of Teacher *B*.

It may be noted also that the coefficients for Teacher *C* in School *Y* are in line with the respective coefficients for the other teachers in School *X*.

The middle set of coefficients in Table 4 is for dominative contacts of the teachers. For the most part these coefficients are identical with, or very close to, the corresponding coefficients for the total number of contacts. In all cases they indicate a high degree of consistency in the observation and recording of this kind of data.

It may be noted again, however, in comparing Teachers *A* and *B* that although *B* in all comparisons had fewer contacts than *A*, the observers were still able to record her dominative contacts with *greater consistency than they could record the dominative contacts* of Teacher *A*. This difference in all probability is to be explained in part by the nature of the program and in part by the difference in techniques of the two teachers. Over half of the contacts of Teacher B were made during the music period with the teacher generally at the piano facing both the children and the observers. Recording was easier. In general, however, Teacher *A* was more subtle in her contacts with children as contrasted with a directness in

Teacher *B*, and *B* usually spoke with more volume than Teacher *A*.

The consistency of recording integrative contacts with individual children as shown in Categories 15-23 inclusive is low in all cases. It can be noted, however, that with one exception the coefficients all fall within a narrow range of nine points, from .44 to .53. The exception is the coefficient of .10 which is based on only 20 tallies and should be excluded from consideration. It can be noted also that the numbers of tallies for integrative contacts range from about one-half to one-third the numbers of dominative contacts and in the case of the coefficient of .10 the domination contacts outnumber the integrative contacts five to one. An examination of the data in this and the previous study (6) together with more qualitative observations of the psychological interplay between teachers and children offers some explanations.

Domination is defined as rigid behavior in which the dominating person knows what he wants. Integrative behavior is flexible behavior in which the person is seeking evolving purposes that take into account the phenomena of change through the confronting of differences. By the nature of the behavior it might be expected that the definiteness of domination would be more easy to recognize than the flexibility of response of one seeking rather a better understanding of another's purposes than the accomplishment of a fixed objective of his own. By its very nature integrative behavior would appear in a wider variety of responses and in more subtle expressions than would domination.

These inherent differences that tend to make integrative behavior more difficult to observe are borne out by the evidence from previous studies. It has been noted (6, pp. 355-358) that preschool teachers were better able to rate children for dominative behavior than for integrative or coöperative behavior. If ratings are based on past observations, this would mean that for these trained preschool teachers integrative behavior was more difficult to observe or to take note of than was domination. The greater predictability in the teachers' ratings of dominative behavior in children was not due to greater frequencies of domination as shown in the experimental situation; for in practically all cases the mean scores for domination were lower than the mean scores for integrative behavior.

Another explanation for low coefficients of consistency of the observers is found in the comparatively lower frequencies of the observations of integrative contacts.

An inspection of the scatter diagrams of these coefficients showed large frequencies in the 1-0 and 0-1 class interval squares. These frequencies indicate that one observer saw a considerable number of contacts which the other observer missed, and vice versa.

However, the data given above show that the two observers obtained virtually an identical total number of observations of teachers' contacts. The inferences are justified that: (*a*) the real *number of teachers' contacts which the observers failed to record* is approximately the same for each of the two observers; and (*b*) some of the teachers' contacts which the two observers missed were not the same contacts, but different ones (as shown on the scatter diagrams in the 0-1, 1-0 class intervals). If the speed of teachers' contacts was greater than the capacity of the observers to observe and record, the observers could have perfect agreement in what they chanced to notice simultaneously and still have low coefficients of consistency because of the numbers which were recorded by one and not by the other. To the extent then that the interplay between the teachers and the children was too swift to be recorded by these methods the low coefficients are not representative of "inconsistency of simultaneous observations" alone. It seems apparent that the low coefficients are to be explained in part by the mechanical difficulties of the method of recording.

3. *Consistency in Recording Individual Categories*

At all times during the study the experimenters have been aware that the individual categories were set up arbitrarily for convenience in recording. Because, however, the question of degree or intensity of *domination* appears to be amenable to measurement and therefore to constitute a problem of importance for future research, the data were analyzed to determine the consistency of the observers in recording teachers' contacts by category. On each observation blank the total number of tallies in each category was computed. These sums were then correlated for the two observers for all categories in which one or more tallies were recorded by one or both observers during a minimum of 25 observation periods. No coefficients were computed for a number less than 25. All coefficients reported are for observations at School *X*; for School *Y* there were only 13 observation periods.

Table 5 gives the coefficients showing the consistency with which

TABLE 5

COEFFICIENTS OF CONSISTENCY OF OBSERVERS FOR CATEGORIES OF TEACHER CONTACTS WITH CHILDREN GROUPED

(One tally on the scatter diagram represents, e.g., Σ tallies for all children for Category 1 on one observation blank. N = number of blanks on which there were, one or more tallies by one or both observers for each respective category.)

Teachers	*A & B*		*A & B*		*A*		*B*		*A & B*	
Session	A.M.		P.M.		A.M. & P.M.		A.M. & P.M.		A.M. & P.M.	
Obs. min.	155.5		187.0		204.5		138.0		342.5	
Mean period	4.44		4.92		4.65		4.76		4.69	
Category	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*
1	32	.92±.02	36	.88±.03	41	.85±.08	27	.96±.01	68	.90±.02
3									25	.96±.01
4					30	.10±.12			54	.25±.09
5	27	.88±.03	28	.63±.08	33	.67±.06			55	.71±.05
7	28	.72±.09	27	.28±.12	34	.80±.05			55	.62±.06
9			27	.65±.08	26	.81±.05			42	.68±.06
15			25	—.27±.13	28	.34±.11			42	—.08±.10
16									39	.81±.04
19			28	.75±.06	33	.88±.03			47	.74±.05
21									34	.27±.11
24	26	.76±.06			34	.66±.07			50	.64±.06

two observers recorded the same number of tallies per category per observation period. In these coefficients the "*Group*" and "*Unidentified*" contacts were treated statistically the same as for individual children. It can be noticed that such a fine analysis leaves blank spaces for several categories. It is obvious that the study of categories of teachers' contacts involves more intensive and extensive research than was contemplated in the present investigation.

All the coefficients in Category 1 (*Determines a detail of activity*), show a high degree of consistency in the observations of the two observers. This category had the highest frequency of tallies of all.

Category 2 (*Direct refusal*) had so few tallies as to make a coefficient of correlation meaningless.

Category 3 (*Relocating or reseating a child*) had tallies recorded in 25 observation periods and yielded the high coefficient of .96 for all the periods both morning and afternoon and for both teachers combined. The observers felt that this sort of contact was relatively easy to identify.

Category 4 (*Postponing, slowing up the child*) was difficult to record. The observers found it difficult in some instances to distinguish between slowing up the fast ones with no apparent disapproval of the activity itself and obstructing activity of which the teacher disapproved, which would be checked in Category 5.

Category 5 (*Disapproval, blame or obstruction*), however, has relatively high coefficients of consistency. The coefficient for both teachers for both sessions is .71; for both teachers for the morning session alone it is .88, and for afternoon alone .63. Teacher *A* for both morning and afternoon shows a coefficient of .67.

Category 6 (*Warning, threats and conditional promises*) was too infrequent to permit correlations.

Category 7 (*Call to attention or to group activity*) showed a high consistency when the contacts of the leading teacher *A* were considered alone. The coefficient of .72 is relatively high for both teachers with the younger children in the morning, but the correlation is very low for both teachers in the afternoon. For all the contacts combined the coefficient of consistency of two observers for this category is .62.

Category 8 (*Rations material*) was too infrequent to permit correlations, as was also Category 10 (*Questions, lecture method*).

Category 9 (*Lecture method*) had low numbers but frequencies in a sufficient number of observation periods to justify three coefficients. The rationing appears to be done chiefly by the head teacher, whose contacts can be observed and recorded by two observers with a fair degree of consistency.

The contacts recorded under Category 15 (*Perfunctory question or statement*) were often very subtle and confusing. The observers were instructed to disregard the perfunctory remarks when they were combined with some contact representing a stronger response and to record only the stronger contact. Some of these perfunctory contacts that were barely non-dominating were found in the "*Yes—but—*" remarks of the teacher. For example, the child said, "*Isn't my basket cute?*" The teacher's response was "*Yes,—but we haven't got it all pasted. The decorations* (which were prepared for the child by the teacher) *are not on.*" The mildly indifferent "*Yes,*" by itself, would have been checked as a perfunctory remark showing that ostensibly the teacher was not in disagreement with the child. The "*Yes*" taken in its immediately following context meant to the observer that the "*Yes*" meant "*No*"*;* that the whole response of the teacher to the child's question was an expression of rigidity in the teacher; that satisfying the child's criteria was not enough to make the basket "*cute*"*;* in order to be cute the basket had to have on it the decorations planned by the teacher.

For Category 16 (*Approval*) the contacts were so infrequent as to yield only one coefficient obtained by combining all the data. The coefficient was .81.

Category 17 (*Accepts difference*) was difficult to record and so low in frequency as not to afford a coefficient of consistency. Category 18 (*Extends invitation to activity*) was also too infrequent to permit correlation of records by two observers.

Category 19 (*Question or statement regarding child's expressed interest or activity*) yields a coefficient of .88 for Teacher *A* alone and of .75 and .74 for data including Teachers *A* and *B* respectively for the afternoon alone and for both morning and afternoon combined.

Category 20 (*The build-up*) or presumably the highly integrative behavior in problem-solving situations had frequencies too small to permit correlations for consistency of observers.

Category 21 (*Participates in joint activity with children*) was so infrequent as to yield only one coefficient for all data combined and that showing no consistency in observations.

Categories 22 (*Sympathy*) and 23 (*Permission*) were too low in frequencies to afford correlations.

Category 24 (*Undetermined*) has a coefficient of .64 for all data combined. Two other correlations based on smaller numbers yielded higher coefficients of .66 and .76 respectively.

In general the analysis of these data by categories is very encouraging. In the first place the data are taken not from experimental situations devised for sampling purposes but from actual schoolroom situations. In the second place the data represent an attempt not to collect samples of certain kinds of contacts but to collect *all* of the observable contacts of all kinds between the teacher and the individual child and between the teacher and the group. The evidence of this study shows that these categories are useful as they stand, but it also justifies further experimentation in the refining of the categories which would produce higher degrees of consistency in recording by individual categories teachers' contacts with children. If the teachers' total observable contacts with children could be broken down into categories sufficiently refined as to yield a generally high measure of consistency of observation it is possible that such data would have high psychological value in understanding and predicting the behavior of children in the schoolroom and also in differentiating teachers for their mental hygiene contribution to individual children and to the group in their ordinary classroom contacts.

4. *Consistency of Recording All the Contacts per Category per Child*

The consistency of observations by two observers has been analyzed and discussed above in terms of total numbers of contacts per observation period, total numbers of combinations of contacts per child per observation period, and total number of contacts with all children per category per observation period. Two further and finer analyses were made. The tallies for each category were correlated child by child for two observers. This required of the observer exactness in identifying both the child and the category. Separate coefficients for each category for each teacher separately for morning and afternoon and for combinations of teachers and sessions were computed. The coefficients were computed not for their present bearing on the study of the general groups of dominative and integrative contacts but to reveal the nature of subsequent research problems. They constitute a large table of 73 coefficients with accompanying probable errors and numbers of cases and are not reproduced here. In general the coefficients except for Category 1 were erratic; in Category 1 they were consistently high.

The same sort of analysis was made for all the data combined instead of computing separate coefficients for each category and for different combinations of teachers and sessions. In obtaining these coefficients, one for School *X* and one for School *Y*, the scatter diagrams were composed of tallies taken from all the observation blanks square by square. This method necessitates exactness in the recording by category and by child each observed contact of the teacher. There is no more rigorous criterion for consistency of observers in recording all the data than the one here employed. Data were recorded from 1,560 squares on record blanks for School *X* and from 378 squares for School *Y*. The coefficients of consistency of the two observers as shown in Table 6 are .78 for School *X* and .77 for School *Y*.

Table 6 shows that the consistency of observers in recording the kind of data sought in this study is sufficiently high to be acceptable as measures of teachers' behavior and that the smaller sampling in School *Y* is practically identical with the large sampling in School *X*.

TABLE 6

COEFFICIENTS OF CONSISTENCY FOR ALL THE DATA OF TWO OBSERVERS IN RECORDING TEACHERS' CONTACTS BY CHILD AND BY CATEGORY

School	N	r	PE_r
X	1,560	.78	.01
Y	378	.77	.02

G. CONSISTENCY OF THE TIME SAMPLINGS

1. *Consistency within a Given Activity*

It was mentioned above that on each observation blank was written the nature of the general activity of the children during that observation period. These designations of activity are necessarily arbitrary and in many particular situations are crude. It was impossible at many times to note a precise moment when one activity stopped and another began. In fact it appeared that in some of the transition periods when the children were finishing one activity and turning to another, the teachers' contacts occurred at highest frequencies. The classification of activities seemed in general too vague to offer comparisons. With two exceptions, however, singing and handwork, there was sufficient continuity and sufficient uniformity of procedure from day to day to offer some comparison. For these activities the data were analyzed to see whether for a given teacher any trends in frequencies were *indicated between consecutive observation periods. Data here are obtained from School X.*

Tables 7 and 8 show respectively for singing activity and handwork the frequencies of teachers' contacts by category during adjacent five-minute periods at School *X* for the morning session. The correlations are based on all possible pairs of observation periods

TABLE 7

SINGING ACTIVITY

Frequencies of teacher's contacts by category during adjacent five-minute periods at School *X*, morning session.

Day	May 20					May 26			May 31		June 3	
Hour	9:59	10:04	10:09	10:14	10:19½	9:48	9:53	9:58	9:51	9:55	9:05	9:10
Category												
1	23	20	21	10	14	28	35	23	13	20	12	6
2	—	—	—	—	—	—	—	—	—	—	—	—
3	—	—	—	—	—	0	1	0	—	—	—	—
4	—	—	—	—	—	0	2	0	1	0	1	0
5	1	7	8	2	1	2	2	2	1	0	2	2
6	—	—	—	—	—	0	2	1	—	—	—	—
7	3	5	5	2	0	0	3	2	4	2	0	2
8	—	—	—	—	—	—	—	—	—	—	—	—
9	0	1	0	0	2	3	3	5	4	1	0	2
10	0	5	1	3	1	3	5	0	5	1	0	1
15	—	0	1	0	1	1	0	1	—	—	1	0
16	1	0	0	1	0	2	2	1	1	1	1	2
17	—	—	—	—	—	—	0	1	0	2	—	—
18	—	—	—	—	—	—	0	4	0	5	1	5
19	0	1	0	—	—	0	1	0	0	1	0	4
20	—	—	—	—	—	—	0	1	—	—	—	—
21	—	—	—	—	—	—	—	—	0	1	0	2
22	—	—	—	—	—	—	—	—	—	—	—	—
23	—	—	—	—	—	—	—	—	0	1	1	3
24	1	0	—	—	—	—	—	—	—	—	1	0

$r = .90.$

TABLE 8
HANDWORK ACTIVITY
Frequencies of teacher's contacts by category during adjacent five-minute periods at School X, morning session.

Day	May 25							May 26	
Hour	9:10	9:15	9:20	9:25	9:30	9:35	9:40	9:15	9:19
Category									
1	13	11	7	12	3	3	9	12	11
2	—	—	0	1	0	—	—	—	—
3	—	—	—	—	—	0	1	—	—
4	0	2	0	2	3	1	3	1	0
5	7	3	2	11	6	6	7	6	7
6	0	2	2	0	0	2	1	—	—
7	2	0	1	1	0	1	1	7	0
8	—	0	1	0	—	—	—	—	—
9	0	6	0	9	2	2	1	0	1
10	1	1	3	0	2	0	—	2	0
15	4	1	0	1	2	0	—	0	2
16	7	8	2	5	6	4	0	3	6
17	—	—	—	—	—	—	—	1	0
18	—	0	8	0	—	—	—	0	1
19	0	3	0	8	9	7	4	14	1
20	—	—	0	1	1	1	0	0	1
21	4	8	3	1	3	3	0	4	2
22	—	—	—	—	—	—	—	—	—
23	—	0	1	0	4	0	1	1	1
24	2	0	1	1	3	3	1	4	5

$r = .41$.

during which there occurred in one or both periods one or more teacher contacts. Zeros were used only when adjacent to recorded observations when they were necessary to constitute a pair. The coefficient for the singing period was .90 and for handwork .41. The high coefficient for the singing period would indicate that no increase or decrease in the number of contacts throughout the singing period may be expected.

The coefficient of .41 does not justify the same statement regarding the handwork period. Another method of analysis taken from Fisher (9, pp. 117 ff.) shows no evidence of a trend upward or downward in number of teachers' contacts. If there existed such a trend the differences between the number of contacts for two consecutive five-minute periods would show a mean value significantly different from zero. When the individual categories are tested by this criterion, the means of the differences either for singing or for handwork are not significantly different from zero in any of the cases for which such testing was justified on the basis of sufficient numbers. Probabilities according to the Fisher method were in all cases greater than .5, which means that deviations of the means from zero could be ascribed to chance.

From the above data it can be said that no systematic tendencies for teachers' contacts to increase or decrease from one five-minute period to the next were found in those observation periods during which the activity could be satisfactorily defined. It may be pointed out that these data justify no statement about a possible cycle within a particular activity, as for example a large number of contacts when beginning a new activity, a small number while the activity is in progress, and a possible increase in the transition to the next activity. This problem would need to be studied as a special project in which there would be a more accurate method of defining an activity, a better control of timing the observation periods, and larger samplings of activities to be analyzed.

2. *The Representativeness of the Time Sampling*

In computing coefficients of consistency of two observers it was found that the scatter diagrams showed not a normal but a skewed distribution of tallies. In fact a high percentage of tallies fell in three squares: 0-1, 1-1, 1-0. The *J*-shaped distribution of these data made it seem desirable that some technique other than the ordinary correlation analysis be employed for determining the consistency of time samples. Accordingly the mean numbers of contacts per hour were computed for the two halves of the total number of observations in all possible groupings of teachers and sessions. On the one hand the contacts per hour for all children combined were employed with categories in rank order for the two time samples. On the other hand the categories were grouped according to types of contacts and the children were arranged in rank order for the two equivalent time samples. It is obvious that by this latter method of using the rank order of children the morning and afternoon groups could not be combined. The differences between the two time samples were then analyzed in two ways: by rank order correlations and by determining the significance of the differences between mean numbers of contacts per hour for the two time samplings.

TABLE 9

DIFFERENCES BETWEEN TWO EQUAL TIME SAMPLINGS OF OBSERVATIONS OF TEACHERS' CONTACTS WITH ALL CHILDREN COMBINED, FOR TEACHERS INDIVIDUALLY AND COMBINED DURING MORNING AND AFTERNOON SESSIONS: PROBABILITIES THAT FOR THE TWO TIME SAMPLINGS DIFFERENCES BETWEEN MEANS OF THE CONTACTS PER HOUR PER CATEGORY ARE DUE TO CHANCE; RANK ORDER CORRELATIONS OF TEACHERS' CONTACTS WITH CHILDREN WITH COEFFICIENTS COMPUTED FROM RANK ORDERS OF CATEGORIES

Teacher	Session	Probability that differences are due to chance	Rank r	Time of obs. (min.)
A & *B*	A.M. & P.M.	$.7 < P < .8$	.97	461
A & *B*	A.M.	$.5 < P < .6$	.91	230
A & *B*	P.M.	$.4 < P < .5$	.97	231
A	A.M. & P.M.	$.8 < P < .9$	.93	337
A	A.M.	$.8 < P < .9$	.88	181
A	P.M.	$.8 < P < .9$	.89	176
B	A.M. & P.M.	$.6 < P < .7$	.86	103
B	A.M.	$P < .01$	.65	49
B	P.M.	$.5 < P < .6$	.80	54
C	A.M.	$.2 < P < .3$	.93	220

The probabilities are computed according to the method reported by Fisher (9, pp. 117 ff.). When $P = .05$ or less, the difference between the means are regarded statistically as representing a real difference greater than zero.

Rank order correlations are corrected to Pearson product moment r's according to tables in Odell (11, pp. 229 ff.).

Figures listed as time of observation are mean times in minutes for the two time samplings. The differences between the two samplings are for the above data less than two minutes. A few children were absent from some of the observations. Frequencies of teachers' contacts with individual children are based on actual time in attendance.

Table 9 shows the treatment of differences between two equal time samplings of observations of teachers' contacts with all children combined, for teachers individually and combined during morning and afternoon sessions; the probabilities, according to the Fisher method (9, pp. 117 ff.), that for the two time sampling differences between the means of the contacts per hour are due to chance; and the rank order correlations of teachers' contacts with children with coefficients computed from rank orders of categories.

In Table 9 there is a probability that the obtained difference is not due to chance in only one comparison which is for an observation

TABLE 10

DIFFERENCES BETWEEN TWO EQUAL TIME SAMPLINGS OF OBSERVATIONS SHOWN FOR GROUPS OF CATEGORIES OF TEACHERS' CONTACTS FOR TEACHERS INDIVIDUALLY AND COMBINED DURING MORNING AND AFTERNOON SESSIONS; PROBABILITIES THAT FOR THE TWO SAMPLINGS DIFFERENCES BETWEEN MEANS OF THE CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN ARE DUE TO CHANCE; RANK ORDER CORRELATIONS OF TEACHERS CONTACTS WITH CHILDREN WITH COEFFICIENTS COMPUTED FROM RANK ORDERS OF CHILDREN

Categories		1-24			1-8			9-10		15-23			Time
			Rank	*S-B*		Rank	*S-B*		Rank		Rank	*S-B*	of obs.
Teacher	Session	Prob.	*r*	*r*	Prob.	*r*	*r*	Prob.	*r*	Prob.	*r*	*r*	(Min.)
A & *B*	A.M.	$.2 < P < .3$	.76	.86	$.2 < P < .3$	.54	.70	$.4 < P < .5$	.37†	$.1 < P < .2$	.70	.82	230
A & *B*	P.M.	$.1 < P < .2$	.64	.78	$.9 < P$	.64	.78	$.1 < P < .2$	.60	$.3 < P < .4$	.39	.56	231
A	A.M.	$.8 < P < .9$	.84	.90	$.1 < P < .2$	.58	.73	$.2 < P < .3$	.59	$.01 < P < .02$*	.72	.84	181
A	P.M.	$.9 < P$	.75	.86	$.4 < P < .5$	.66	.80	$.9 < P$	.53	$.8 < P < .9$	.51	.68	176
B	A.M.	$.6 < P < .7$	.05†		$.3 < P < .4$	.09†		$.6 < P < .7$	.35†	$.01 < P < .02$*	.26†		49
B	P.M.	$.05 < P < .1$	—.10†		$.8 < P < .9$	—.15†		$.05 < P < .1$	—.14†	$P < .01$*	—.63†		54
C	A.M.	$.4 < P < .5$	.62	.77	$.6 < P < .7$	.49	.64	$.2 < P < .3$	—.24†	$.2 < P < .3$	.73	.84	220

The probabilities are computed according to the method reported by Fisher (9, pp. 117 ff.). When $P = .05$ *or* less, the differences between the means are regarded statistically as representing a real difference greater than zero.

Rank order correlations are corrected to Pearson product moment *r*'s according to Tables in Odell (11, pp. 229 ff.).

Figures listed as time of observation are mean times in minutes for each of the two time samplings. The differences between the two samplings are for the above data less than two minutes. A few children were absent for some of the observations. Frequencies of teachers' contacts with individual children are based on actual time in attendance.

*Three differences regarded statistically as representing a *real* difference not due to chance.

†Ten lowest coefficients of correlation.

time totaling 49 minutes. This period shows also the lowest coefficient of correlation in the table. It is evident that 49 minutes is too short a time to be representative of the teachers' contacts as recorded by individual categories. It can be seen, however, that 103 minutes appears to show a fair consistency between the two time samples and that anything over 200 minutes may be regarded as highly satisfactory for the data treated in this way.

Table 10 shows the data in these two time samples analyzed in a way different from that employed for Table 9 and with somewhat different results. Table 10 shows the treatment of differences between two equal time samples of observations shown for groups of categories of teachers' contacts (1-24, 1-8, 9-10, 15-23) for teachers individually and combined for morning and afternoon sessions; the probabilities that for the two samplings, differences between means of the contacts per hour with individual children are due to chance; the rank order correlations of teachers' contacts with children, with coefficients computed from rank orders of children.

As compared with Table 9 the lower coefficients in Table 10 are to be explained in part by the mental set of the observers and their methods of recording. The observers were attempting to record contacts that at times occurred so rapidly as to tax their capacities for speed of recording. They watched the teacher with a determination to identify instantly the nature of the teacher's contact. Their attention was on the technique of the teacher. The preceding analyses of the data in this study showed that when the individual children involved in the contacts are disregarded, the observers could record with higher consistency. Tables 9 and 10 show that for the same purposes of recording, i.e., with attention to the individual categories and with the children grouped in the analysis of the data, representative samples of teachers' contact can be obtained in shorter time.

The observers report that in the mental processes of observing and recording it was the identifying of the contact which came first; the identifying of the child with whom the teacher had the contact was a separate and secondary mental step. The fact that the children moved about a great deal while the observers were generally keeping their attention on the teacher explains in part the differences between Tables 9 and 10.

In Table 10 there are three comparisons in which obtained differences between the time samplings according to the Fisher method (9, pp. 117 ff.) are not due to chance. They are all found in the recording of integrative behavior of Teacher *A* in the morning session and of Teacher *B* in both morning and afternoon sessions. None of the other comparisons are found by the Fisher method to represent significant differences. The coefficients of correlation, however, are lower than those in Table 9.

In Table 10 the ten lowest coefficients of correlation are marked. Eight of these 10 lowest coefficients occur for Teacher *B* who was observed for a total of less than an hour in each session. It is obvious that an hour's observation time does not secure a representative sample of teachers' contacts with individual children even when the contacts are grouped as in Table 10.

Of the 10 lowest coefficients the two others are found in the combination of Categories 9 and 10, where the numbers are very small, a grouping which for low frequency and ambiguity has in the present study been excluded from the domination and integration categories, though included in all treatment of the total number of teachers' contacts.

It can be noted that with one exception, unimportant in the present study (for Categories 9-10), the time sampling for Teacher *A* shows higher coefficients of correlation than do the samplings of Teachers *A* and *B* combined, even though the combination of teachers shows higher totals of observation time.

The Spearman-Brown prophecy formula (10, p. 271) was applied to each coefficient of correlation in Categories 1-24, 1-8, and 15-23 with the exclusion of the coefficients for Teacher *B*, for whom the time sampling was obviously too short. These Spearman-Brown coefficients, given in Table 10, represent the prophecy of consistency if the length of time sampling were doubled. Two-thirds of the 15 coefficients are above .75 and with two exceptions they are all .68 or above. In the case of these two exceptions the difference in the original time samplings used was not significantly different by the Fisher method.

In Table 10 the combined data for Teachers *A* and *B* for total contacts in the afternoon session yielded a coefficient of .64. This coefficient was obtained by a rank order correlation of one period of 231 minutes of observation with a second and subsequent period of equal duration. Raised by the Spearman-Brown formula to the equivalent of the total time actually observed the coefficient of .64 becomes .78. The question arose whether the consistency of the period of observation would be expressed differently by split-half correlations of alternate five-minute observation periods. Such a trial correlation was made on these data because they included the total contacts of both teachers for the longest observation time represented in Table 10. The split-half coefficient of correlation was .61 which when raised by the Spearman-Brown formula to the time actually observed became .76.

From Table 10 it can be concluded that measuring the teachers' contacts with individual children cannot show representative frequencies in an observation time shorter than 200 minutes. The children in the morning and afternoon groups were not the same children, and for that reason the data for these two groups cannot be combined.

It is thus possible to say for the comparisons in Table 10 that in general a total observation time of 300 or of 400 minutes is sufficient to afford a representative sample of the teachers' contacts with individual children. It should be pointed out that the interpretations of Tables 9 and 10 apply only to time samples of data gathered over a few consecutive days.

3. *The Representativeness of Two Twenty-Minute Periods of the Same Activity*

Among the observation data there were found two sets of data, each recorded during twenty minutes of handwork for contacts of Teacher *B* on consecutive mornings. The correlation of the teacher's contacts for these two periods yielded a coefficient of .94.

H. DIFFERENCES BETWEEN TEACHERS *A* AND *B*

Some of the differences between Teachers *A* and *B* have been pointed out. Computations have been made to test the significance of some of the observed differences.

1. *Differences Shown by Consistency of Observers: Categories 1-8, Table 4*

It was pointed out above that Table 4 showed that in all combinations of data in which Teachers *A* and *B* were treated separately the observers were able to record contacts in Categories 1-8 more consistently for Teacher *B* than for Teacher *A*. The three pairs of coefficients of consistency for Teachers *A* and *B* in the morning and afternoon sessions separately and combined were tested by the Fisher method for the significance of the obtained differences. Table 11 gives the coefficients, the differences, *Z*-score differences, standard error of the *Z*-scores, and the critical ratios. All of the differences are found to be significant.

TABLE 11

DIFFERENCES BETWEEN TEACHERS *A* AND *B* AS SHOWN BY THE FISHER METHOD OF TESTING COEFFICIENTS OF CORRELATION OF CONSISTENCY OF TWO OBSERVERS FOR DOMINATION CONTACTS, CATEGORIES 1-8

(Coefficients are taken from Table 4)

Session	Coefficients *A*	*B*	*Diff.*	*Z*-score *diff.*	*SE* of *Z*-score	$\frac{Z\text{-score}^*}{SE_{Z\text{-score}}}$
A.M. and P.M.	.89	.96	.07	.5240	±.0781	6.71
A.M.	.87	.97	.10	.7592	±.1271	5.97
P.M.	.80	.94	.14	.6396	±.1142	5.60

*Critical ratios greater than 2 are considered to indicate differences between coefficients that are statistically significant.

2. *Differences Shown by Consistency of Observers: Category 1, Table 5*

Referring again to Table 5 it can be seen that for Category 1 Observers *M* and *N* had coefficients of consistency of .85 for Teacher *A* and .96 for Teacher *B* for the combined morning and afternoon sessions. Using the method suggested by Fisher (9, pp. 190 ff.) it was found that these differences were significant. Since Category 1 had considerably more tallies than any other category it is seen that there is a methodological problem involved when observing different teachers, or when observing different situations. Since over half of the observation time for Teacher *B* occurred during the music period the question is still open as to what extent the nature of this activity made it easier to record the kind of contacts classified as Category 1. It is clear that much more comprehensive studies are indicated, involving observations on a number of teachers in corresponding schoolroom activities.

3. *Differences Shown in Rank Order of Children*

Children in the morning and afternoon sessions of School *X* were grouped in rank order of number of contacts per hour with Teachers *A* and *B* respectively. Separate rank orders were made for the first half of the total observation time, the second half of the total observatime, and for the total observation time. Table 12 gives the rank order correlations respectively for the several groups of contacts corrected to Pearson product-moment coefficients from the table in Odell (11, pp. 229 ff.). The coefficients in Table 12 indicate that within the time samplings available for this analysis Teachers *A* and *B* are different in respect to the number of contacts per hour per child.

TABLE 12

DIFFERENCES BETWEEN TEACHERS *A* AND *B* AS SHOWN IN RANK ORDER CORRELATIONS OF CHILDREN RANKED ACCORDING TO FREQUENCY OF CONTACTS PER HOUR WITH TEACHERS *A* AND *B* RESPECTIVELY FOR THE FIRST AND SECOND HALVES RESPECTIVELY OF THE TOTAL OBSERVATION TIME AND FOR THE TOTAL OBSERVATION TIME

Rank order coefficients are corrected to Pearson product moment *r*'s from the table in Odell (11, p. 229 ff.).

	Observation time					
	A.M.			P.M.		
Categories	Total	First half	Second half	Total	First half	Second half
1-24	.47	.29	.52	.18	.13	.39
1- 8	.43	.30	.73	.14	.19	.48
9-10	.14	.51	.06	.22	—.35	.09
15-23	.28	.56	.11	.19	.24	.39

I. DIFFERENCES IN CONTACTS DURING THE MORNING AND AFTERNOON FOR THE SAME TEACHERS

By combining the data for Teachers *A* and *B* there were sufficient data in Category 1 to make comparisons between teachers' contacts during the morning and afternoon sessions. When these data were analyzed according to the Fisher method (9, pp. 190 ff.) there were found to be only 79 chances in 100 that the obtained difference represented a real difference greater than zero. These data, however, combine two teachers who are known to be different in their contacts with children. By combining all the domination contacts (Categories 1-8) there were sufficient data to compare each teacher with herself. The Fisher method of Z-score comparison of mean

TABLE 13

Differences between Teachers' Contacts in the Morning and Afternoon for Teachers *A* and *B* Respectively as Shown by Testing the Coefficients of Correlation of Consistency of Two Observers for Domination Contacts, Categories 1-8, by the Fisher Method

Coefficients are taken from Table 4.

Teacher	Coefficients			Z-score	*SE* of	$\frac{Z\text{-score}^*}{SE_{Z\text{-score}}}$
	A.M.	P.M.	*Diff.*	*diff.*	Z-score	
A	.87	.80	.07	.2346	±.1133	2.07
B	.97	.94	.03	.3542	±.1279	2.77

*Critical ratios greater than 2 are considered to indicate differences between coefficients that are statistically significant.

scores (9, pp. 190 ff.) yielded the results shown in Table 13. Each teacher in her contacts with respective morning and afternoon groups of children is seen to show differences that are statistically significant.

J. Number of Teachers' Contacts per Hour

1. *Number of Teachers' Contacts per Hour per Category and for Combinations of Categories*

It was mentioned previously that in an early stage of this study it was thought that an experimental situation might be necessary in order to economize on the experimenters' time. It was discovered, however, that the teachers' contacts with children at times happened with such rapidity as to tax the capacity of the observers for speed of recording. The data consistently show that in all probability the observers were unable to record contacts at the fastest speed with which they occurred and that this limitation lowers the coefficients of consistency. To the extent that the coefficients of consistency are thus lowered they reflect a lack of simultaneity in observations, a discrepancy that can exist apart from inconsistency in the abilities of the two observers to observe and record.

The speed or rapidity with which teachers make contacts with children raises some pedagogical and mental hygiene questions. The complaint has often been made at home and at school that children are unable to concentrate, that they cannot carry on activities by themselves or hold to a given purpose without adult encouragement or stimulation. The complaint though frequently made is not very clearly formulated. Studies have been made of the attention span of preschool children and of others, but as yet there are no standards or criteria against which to evaluate either the performance of an individual child or that of a group. From experience in clinical psychology one has often suspected that children have been unnecessarily interrupted in their serious purposes by well-meaning adults. In some cases the "over-supervision" has been so unrelenting as to make it seem as though the child could do little or nothing by himself. In fact the greater amount of "free play" or freedom to inquire and to explore one's environment is one of the chief criteria by which the nursery school is distinguished from public school education.

But as to how much free play a child needs; how much supervision is a "good thing"; how many contacts with an adult a child should have; when supervision ceases and "over-supervision" begins; what mistakes and how many a child should be permitted to make without adult interference;—for answers to these questions there are only unreliable clinical generalizations. It is obvious that before one can speak reliably about "too much" he must first have units of measurement. This study constitutes an important first step in providing such units of measurement.

Table 14 shows the mean number of contacts per hour which each teacher had with individual children; the mean numbers of domination and integration contacts per hour, and the mean number of contacts per hour classified according to individual categories. Table 14 gives also the total time in minutes on which are based the respective rates per hour of contacts by individual teachers according to session and to combinations of teachers and sessions. The Domination-Integration ratio is obtained by dividing the mean number of domination contacts per hour by the mean number of integration contacts per hour.

It can be noted that the highest frequencies for all contacts for separate periods show for Teacher *A*, 421.3 contacts per hour; for Teacher *B*, 474.5 contacts per hour; and for Teacher *C*, 489.8 contacts per hour. When morning and afternoon contacts are added together and the means per hour per child computed, Teachers *A* and *B* are nearer together in frequencies, showing respectively 401.3 and 431.7 contacts per hour. This represents for Teacher *A* a rate of 6.7 contacts a minute for 11.9 hours of observation, and for Teacher *B* a rate of 7.2 contacts per minute over a period of 3.5 hours of

TABLE 14

MEAN NUMBER OF CONTACTS PER TEACHER PER HOUR FOR EACH CATEGORY AND FOR COMBINATIONS OF CATEGORIES

School	*X*									*Y*
Teacher	*A*			*B*			*A* and *B*			*C*
	A.M.	P.M.	A.M. and P.M.	A.M.	P.M.	A.M. and P.M.	A.M.	P.M.	A.M. and P.M.	A.M.
Min. obs.	362.0	352.5	714.5	98.5	109.0	207.5	460.5	461.5	922.0	440.5
Category										
1-24	421.3	380.8	401.3	384.4	474.5	431.7	413.4	402.9	408.4	489.8
1- 8	244.2	210.4	227.5	278.4	350.6	316.3	251.5	243.5	247.6	292.4
9-10	46.6	57.9	52.2	46.9	51.7	49.5	46.6	56.4	51.6	63.8
15-23	115.7	97.5	106.9	55.4	71.6	63.9	102.8	91.4	97.1	126.3
1	143.4	128.0	135.8	202.8	254.9	230.2	156.1	158.0	157.1	167.0
2	1.8	1.0	1.4	.0	.6	.3	1.4	.9	1.2	1.0
3	4.6	3.4	4.0	4.3	1.1	2.6	4.6	2.9	3.7	3.7
4	13.6	15.5	14.5	3.7	7.2	5.5	11.5	13.5	12.5	12.4
5	48.9	38.1	43.6	25.0	35.2	30.4	43.8	37.4	40.6	64.1
6	5.0	5.3	5.1	3.1	2.2	2.6	4.6	4.6	4.6	6.4
7	21.2	14.5	17.9	22.5	35.8	29.5	21.5	19.5	20.5	29.7
8	5.6	4.6	5.1	17.1	13.8	15.3	8.1	6.8	7.4	7.9
9	39.3	36.6	38.0	24.4	25.9	25.2	36.1	34.1	35.1	58.8
10	7.3	21.3	14.2	22.5	25.9	24.3	10.6	22.4	16.5	4.9
15	18.6	15.2	16.9	7.9	6.1	6.9	16.3	13.0	14.7	18.8
16	16.7	14.5	15.6	17.1	29.2	23.4	16.8	17.9	17.4	10.4
17	.7	1.5	1.1	1.8	1.7	1.7	.9	1.6	1.2	1.8
18	8.3	6.3	7.3	11.0	3.9	7.2	8.9	5.7	7.3	4.0
19	34.6	22.6	28.7	6.1	14.9	10.7	28.5	20.8	24.7	41.0
20	13.1	11.1	12.1	3.7	8.3	6.1	11.1	10.4	10.7	8.2
21	19.7	24.9	22.3	4.9	6.1	5.5	16.6	20.4	18.5	38.7
22	.3	.2	.3	.0	.0	.0	.3	.1	.2	1.8
23	3.7	1.4	2.5	3.1	1.7	2.3	3.5	1.4	2.5	1.8
24	14.9	15.0	15.0	3.7	.6	2.0	12.5	11.6	12.1	7.4
D-I ratio	2.1	2.2	2.1	5.0	4.9	4.9	2.4	2.7	2.5	2.3

observation, with the presumption in both cases that a considerable though undetermined number of contacts were unrecorded.

Another pedagogical as well as psychological problem is presented in the comparison of the teacher in School *Y* with the teachers in School X. There were 23 children enrolled in the morning session of School *X* and 21 enrolled for the afternoon. In School *Y*, however, there were only 11 children present during the period of observation for this study. It can be seen in Table 14 that for Teacher *C* there are no great proportional divergences in frequencies of contacts as compared with the frequencies for Teachers *A* and *B* who had larger groups of children. It can be noted that the rates of integrative contacts and of total numbers of contacts of Teacher *C* exceed in all cases the rates for the teachers in School *X* while the rate of dominative contacts of Teacher *C* exceeds all rates except for Teacher *B* for the afternoon and for the combination of morning and afternoon.

Is one to draw an inference that teachers are themselves responding at a "capacity rate" whether they have one dozen or two dozen in the room? Or stated in another way, do teachers regardless of numbers of children before them respond at a fairly constant rate? These data raise a further question as to how many children constitute a teacher load. Were the children in School *Y* receiving twice as much "teaching" as the children in School *X*? The data to be presented below in Tables 15 to 19 inclusive show that the children in School *Y* had almost twice as many individual contacts per hour as did the children in School *X*. Again the question as to how many contacts are desirable pedagogically and psychologically must remain unanswered. All that can be said is that in this study measures have been developed that indicate considerable differences in teachers' techniques.

It may be noted in Table 14 that in all combinations of data the dominative contacts (Categories 1-8) outnumber the integrative contacts (Categories 15-23) by at least two to one, and that in one case the ratio of domination to integration contacts is five to one. The contrast between teachers' dominative contacts and integrative contacts is shown graphically in Figure 2 which gives the respective frequencies per hour for individual teachers. The data in Table 14 may be said to constitute a measure of one aspect of the "psychological teaching load" of these kindergarten teachers, or of what it means in terms of dominative and integrative contacts with individual children to be a teacher in these schoolrooms.

2. *Number of Teachers' Contacts per Hour per Individual Child for Groups of Categories*

Five different groups of data have been analyzed to show the mean number of contacts per hour per child for Teachers *A* and *B*

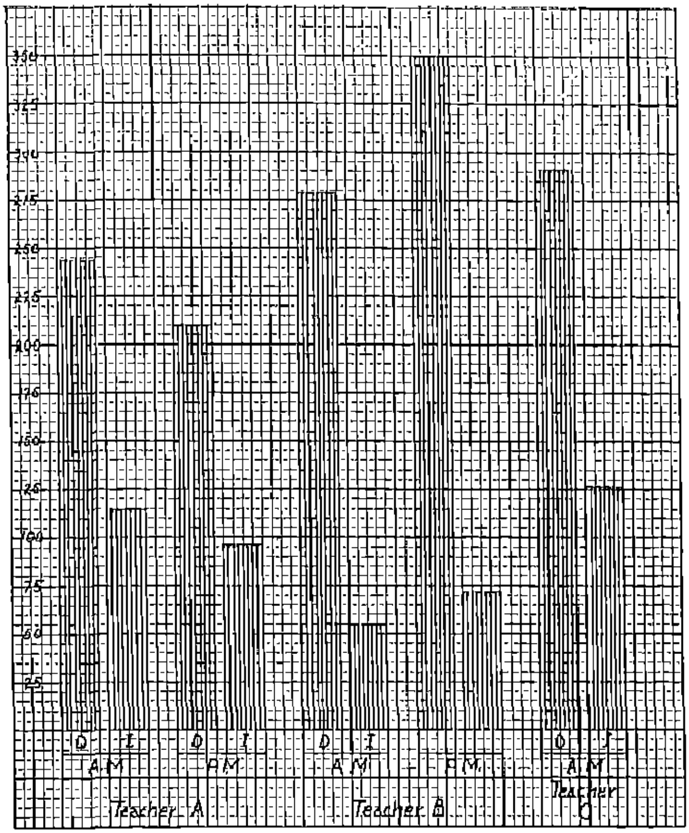


FIGURE 2

Mean Number of Dominative and Integrative Contacts per Hour Which Teachers *A*, *B*, and *C* Respectively Had with Individual Children (from Table 14)

individually with children enrolled respectively in morning and afternoon sessions of School *X*, and for Teacher *C* whose children were enrolled only for a morning session in School *Y*.

Tables 15 to 19 inclusive show for each child the mean number of total contacts per hour, the mean number of dominative contacts per hour, and the mean number of integrative contacts per hour, together with the respective rank orders of children and the *D-I* ratios. The corresponding data are presented graphically in Figures 3 to 7 inclusive.

a. Teacher A, morning session. Table 15 and Figure 3 show the mean numbers of contacts per hour which Teacher *A* had with individual children enrolled in the morning session of School *X*. The children are listed in rank order of mean total numbers of teacher's contacts per hour, Categories 1-24. From Figure 3 it can be seen at a glance that as far as can be indicated by frequency of teacher's contacts per hour the individual children in this kindergarten live at school in different environments.

TABLE 15

MEAN NUMBER OF TEACHER'S CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN; TEACHER *A*, MORNING SESSION

Child	Total contacts Mean *N* per hr.	Total contacts Rank	Domination Mean *N* per hr.	Domination Rank	Integration Mean *N* per hr	Integration Rank	*D-I* ratio	*D-I* rank
8*M*	39.3	1	24.9	1	8.0	5	3.1	6
3*M*	27.9	2	17.7	2	6.5	8	2.7	7.5
1*F*	23.4	3	12.4	3	8.3	4	1.5	14
20*F*	21.0	4	6.4	14	10.7	1	.6	22.5
22*F*	20.2	5	12.1	4	4.5	11.5	2.7	7.5
13*M*	19.9	6	9.3	5	8.3	3	1.1	18.5
10*M*	18.7	7	7.1	11	9.8	2	.7	21
4*F*	15.4	8	6.5	12.5	7.0	7	.9	20
2*F*	14.9	9	7.3	10	5.2	9	1.4	15.5
19*F*	14.5	10	9.1	6	3.9	14	2.3	11
16*F*	13.3	11	4.3	21	7.4	6	.6	22.5
18*F*	13.2	12	7.8	8.5	3.4	16	2.3	11
11*M*	12.9	13	6.5	12.5	4.5	11.5	1.4	15.5
23*M*	12.4	14	8.6	7	2.2	18.5	3.9	2
9*F*	12.3	15	7.8	8.5	3.5	15	2.2	13
7*M*	11.9	16	6.1	15	5.1	10	1.2	17
14*F*	10.1	17	4.5	20	4.1	13	1.1	18.5
15*M*	9.1	18	5.7	17	2.2	18.5	2.6	9
12*F*	9.0	19	5.8	16	2.5	17	2.3	11
6*M*	8.2	20	5.4	18	1.7	20	3.2	4.5
17*F*	6.7	21	5.3	19	1.4	21	3.8	3
21*F*	5.0	22	3.2	22.5	1.0	22	3.2	4.5
5*M*	4.1	23	3.2	22.5	.7	23	4.6	1

The range of total number of contacts is from 4.1 to 39.3 per hour, with the median child receiving 13.2 contacts per hour. The median child thus receives about three times as many contacts

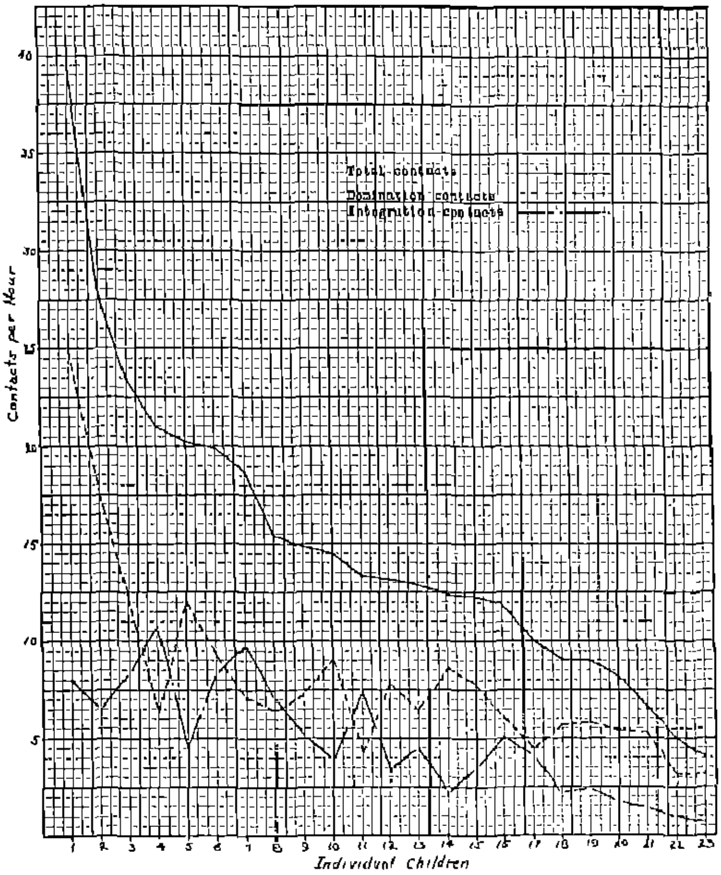


FIGURE 3
Mean Number of Contacts per Hour Which Teacher *A* Had with Individual Children Enrolled in the Morning Session (from Table 15)

per hour as the child lowest in rank and only about one-third the frequency of that of the child highest in rank order. Child 8*M* at the top of the rank order received almost 50 per cent more contacts than Child 3*M* who was next in rank.

The frequencies of dominative contacts show a range of from 3.2 to 24.9, with the median at 6.5 dominative contacts per hour.

The rank orders show generally small differences from child to child from the bottom of the list up to the fifth ranking child who is only 50 per cent above the median. The fourth in rank, however, is about 50 per cent above the fifth in rank or almost twice the median in teacher's dominative contacts.

The range of integration frequencies is from 0.7 to 10.7 contacts per hour, with the median at 4.5. From Figure 3 it can be seen that the curve for integrative contacts not only extends within a shorter range than the curve for domination contacts but that, excepting the cases of three individual children, the integration curve lies below the domination curve. For these three children the Domination-Integration ratios become less than one.

The Domination-Integration ratios range from 0.6 to 4.6. It can be noticed that the child with highest *D-I* ratio is the child who had the lowest total number of contacts per hour with the teacher. In the anticipation of further research it may be asked: is this child to be regarded as "pedagogically self-sufficient" or merely neglected? Or what does it mean to a kindergarten boy whose frequency of total contacts with the teacher is relatively "negligible" to have four out of five of those contacts of a dominative character?

At the other end of the rank order of total contacts is Child 8*M* with a *D-I* ratio of 3.1. What does it mean to a kindergarten boy, who has three times as many contacts with the teacher as the median child, to have three out of four of those contacts of a dominative nature?

One child (12*F*) in the morning group had the highest chronological age (81 months) and the lowest *IQ* (70). She was eight months older than the next oldest child in her group. The influence of this child's presence on means and ranges of chronological age, mental age, and *IQ* for the group was indicated in Table 2. One may wonder what a teacher would do with such a child in a kindergarten group and what interpretation could be made from the measures of teachers' behavior developed in this study.

In Table 15 it may be noted that Child 12*F* is below the median in rank in frequency of teacher's contacts per hour for dominative and integrative contacts and also for total contacts. She is 16th in rank in a group of 23 children in dominative contacts, 17th in rank in integrative contacts, and 19th in rank in frequency per hour of total contacts with the teacher. These ranks indicate that in comparison with the group in general Teacher *A* had relatively few contacts of any kind with this child.

b. Teacher B, morning session. Table 16 and Figure 4 show for the same children as in Table 15 and Figure 3 the mean numbers of contacts per hour with Teacher *B*. The range of total number of contacts is from 3.1 to 34.1 with the median at 7.8, which is only about two-thirds the median for Teacher *A*. The highest ranking child has almost three times as many contacts as the next in rank. From the second ranking child to the lowest child the curve for total contacts in Figure 4 constitutes virtually a straight line that is not very discriminating from one child to the next. The

TABLE 16

MEAN NUMBER OF TEACHER'S CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN; TEACHER *B*, MORNING SESSION

Child	Total contacts Mean *N* per hr.	Rank	Domination Mean *N* per hr.	Rank	Integration Mean *N* per hr.	Rank	*D-I* ratio	*D-I* rank
8*M*	34.1	1	20.7	1	9.1	1	2.3	18
17*F*	13.1	2	7.4	4	3.3	3	2.2	19
1*F*	12.0	3	9.2	3	1.4	12	6.6	10
20*F*	10.9	4	6.7	7.5	3.4	2	2.0	21
3*M*	10.4	5	10.4	2	.0	20.5	∞	3.5
14*F*	9.8	6	7.3	5	2.4	6	3.0	14
16*F*	9.1	7.5	4.3	18	3.1	4.5	1.4	23
23*M*	9.1	7.5	6.1	10.5	1.8	9	3.4	12.5
6*M*	8.5	9.5	6.7	7.5	.6	16	11.2	7
10*M*	8.5	9.5	6.1	10.5	1.8	9	3.4	12.5
9*F*	7.9	11	4.9	15	1.8	9	2.7	16.5
19*F*	7.8	12	4.7	17	3.1	4.5	1.5	22
2*F*	7.3	14	6.1	10.5	.6	16	10.2	8
4*F*	7.3	14	5.5	13	.6	16	9.2	9
12*F*	7.3	14	4.9	15	1.8	9	2.7	16.5
15*M*	6.8	16	6.8	6	.0	20.5	∞	3.5
22*F*	6.7	17	6.1	10.5	.0	20.5	∞	3.5
18*F*	6.1	19	3.7	20	.0	20.5	∞	3.5
11*M*	6.1	19	3.7	20	1.8	9	2.1	20
13*M*	6.1	19	4.9	15	1.2	13.5	4.1	11
7*M*	4.6	21	3.5	22	1.2	13.5	2.9	15
5*M*	3.7	22	3.7	20	.0	20.5	∞	3.5
21*F*	3.1	23	3.1	23	.0	20.5	∞	3.5

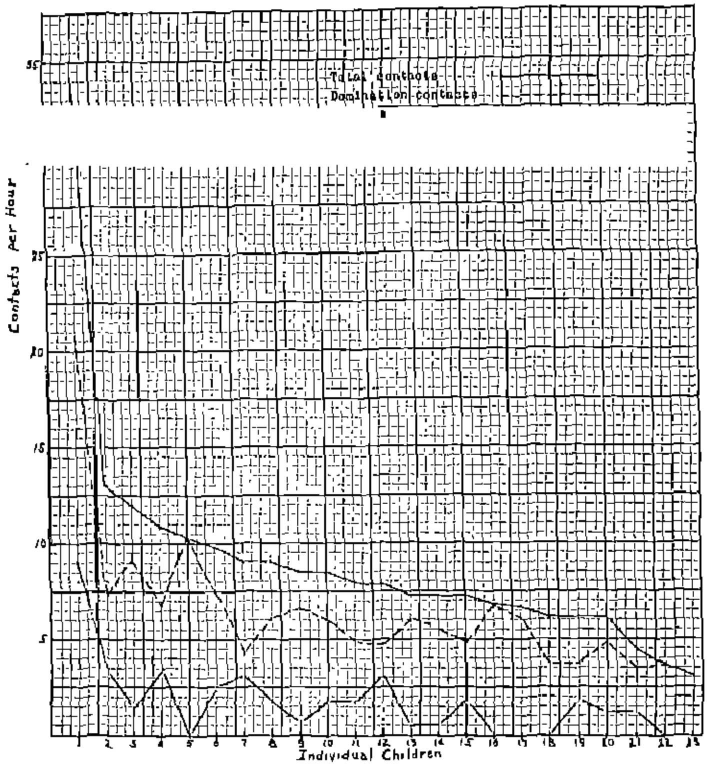


FIGURE 4

MEAN NUMBER OF CONTACTS PER HOUR WHICH TEACHER *B* HAD WITH INDIVIDUAL CHILDREN ENROLLED IN THE MORNING SESSION (FROM TABLE 16)

second child in rank receives, however, four times as many contacts per hour as the lowest child in rank.

The domination contacts range in frequency from 3.1 to 20.7, with the median at 6.1. Again, the first in rank is much above the second, having twice as many domination contacts per hour. It can be seen that the child second in rank in domination contacts received only domination contacts from Teacher *B*, but these were sufficient in frequency to make him rank fifth in the group for total contacts with Teacher *B*.

The curve for integration contacts, *excepting for the child highest in rank*, moves through a narrow range. The range is from zero to 9.1. The median child received 1.4 integrative contacts per hour of observation. Six children received no integration contacts and three others received less than one per hour.

The *D-I* ratios range from 1.4 to infinity, with the median at 3.4. All children received 3.1 or more domination contacts per hour, but six received no integration contacts from Teacher *B*.

It was pointed out above that Child 12*F*, who was the oldest in the group and who had the lowest *IQ*, had relative to the group few contacts with Teacher *A*. Table 16 shows the frequency of contacts which this child received from Teacher *B*. As compared with the frequency of contacts which other children received from Teacher *B*, Child 12*F* was below the median in both dominative and total contacts but above the median in integrative contacts. In interpret-

TABLE 17

MEAN NUMBER OF TEACHER'S CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN; TEACHER *A*, AFTERNOON SESSION

	Total contacts		*Domination*		*Integration*			
Child	Mean *N* per hr.	Rank	Mean *N* per hr.	Rank	Mean *N* per hr.	Rank	*D-I* ratio	*D-I* rank
27*M*	39.7	1	23.2	1	8.2	2.5	2.8	5
28*F*	26.4	2	10.6	3	7.3	5	1.5	12
25*M*	23.7	3	14.7	2	7.4	4	2.0	10.5
43*F*	21.3	4	7.8	7	8.2	2.5	1.0	19
37*F*	18.6	5	10.0	5	4.3	9	2.3	9
35*F*	18.2	6	7.7	8.5	3.2	15	2.4	8
24*F*	16.9	7.5	8.3	6	6.2	7	1.3	15.5
30*M*	16.9	7.5	6.1	13	8.6	1	.7	21
42*M*	15.7	9.5	7.0	10	6.8	6	1.0	19
32*M*	15.7	9.5	7.7	8.5	3.9	13	2.0	10.5
36*M*	15.3	11	10.4	4	1.9	17	5.5	1
38*F*	15.2	12	6.4	12	6.1	8	1.0	19
29*F*	11.1	13	5.5	15	3.9	13	1.4	13.5
41*F*	10.9	14.5	6.5	11	2.6	16	2.5	7
40*F*	10.9	14.5	5.7	14	4.0	11	1.4	13.5
26*M*	9.5	16.5	4.1	17	3.9	13	1.1	17
31*M*	9.5	16.5	5.3	16	4.1	10	1.3	15.5
39*F*	5.8	18	3.7	20	.9	20.5	4.1	2
34*F*	5.6	19	3.9	18	1.5	18	2.6	6
44*F*	5.3	20	3.8	19	1.1	19	3.5	4
33*M*	5.1	21	3.4	21	.9	20.5	3.8	3

ing these ranks, however, it must be remembered that Teacher *B* had relatively few integrative contacts with the children; with six children she had no integrative contacts. Although Child 12*F* ranks ninth (two places above the median) in integrative contacts with Teacher *B*, she had actually fewer integrative contacts per hour with Teacher *B* than she had with Teacher *A*, with whom she ranked five places below the median.

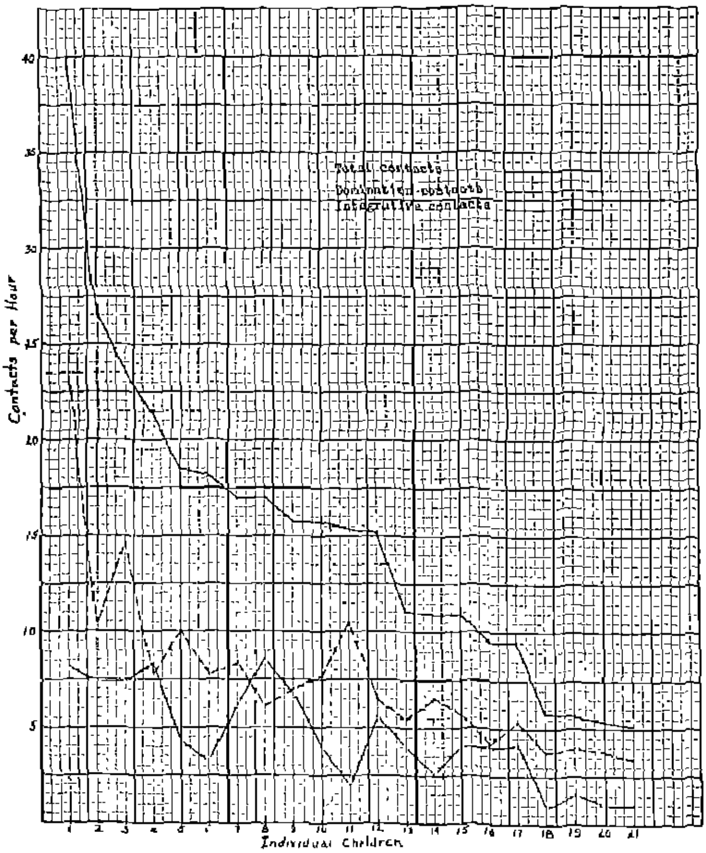


FIGURE 5

MEAN NUMBER OF CONTACTS PER HOUR WHICH TEACHER *A* HAD WITH INDIVIDUAL CHILDREN ENROLLED IN THE AFTERNOON SESSION (FROM TABLE 17) *c. Teacher A, afternoon session.* Table 17 and Figure 5 show the mean number of contacts per hour which Teacher *A* had with the group of children enrolled in the afternoon session. A comparison of Figure 5 with Figure 3 shows a very similar pattern of individual contacts for Teacher *A* in respective morning and afternoon groups of children. In the afternoon group as in the morning group there are two children who are outstanding above the others in total number of contacts per hour. The range of total contacts is from 5.1 to 39.7, almost the same as in the morning group, with the median of 15.3 within two points of the median of the morning group.

For domination contacts the range of from 3.4 to 23.2, and the median of 6.5 are virtually identical with the range and median of the morning group with the same teacher.

The integration contacts range from 0.9 to 8.6, with the median at 4.0.

TABLE 18

MEAN NUMBER OF TEACHER'S CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN; TEACHER *B*, AFTERNOON SESSION

	Total contacts		Domination		Integration			
Child	Mean *N* per hr.	Rank	Mean *N* per hr.	Rank	Mean *N* per hr.	Rank	*D-I* ratio	*D-I* rank
27*M*	35.2	1	22.0	1	5.0	4	4.4	12
35*F*	25.3	2	16.5	2	1.7	14.5	9.7	5
30*M*	25.1	3	15.7	4	8.7	2	1.8	18
32*M*	22.6	4	16.0	3	3.3	6	4.8	10.5
28*F*	19.3	5	9.4	13	9.4	1	1.0	21
29*F*	18.2	6	10.5	11	7.2	3	1.5	19
41*F*	16.0	7	11.6	7	2.8	8	4.1	14
36*M*	15.7	8	12.1	6	2.1	11.5	5.8	7
33*M*	14.3	9	10.5	11	2.2	10	4.8	10.5
34*F*	13.8	10	10.5	11	2.8	8	3.8	15
44*F*	12.9	11	12.9	5	.0	20.5	∞	1.5
31*M*	12.1	12.5	11.0	8.5	.6	18	18.3	3.5
43*F*	12.1	12.5	11.0	8.5	.6	18	18.3	3.5
38*F*	11.9	14	9.1	14	2.1	11.5	4.3	13
37*F*	11.0	15	8.8	15	1.7	14.5	5.2	8
25*M*	10.0	16	7.1	16.5	1.4	16	5.1	9
24*F*	9.7	17	5.6	19	4.0	5	1.4	20
42*M*	9.0	18	7.1	16.5	1.9	13	3.7	16
26*M*	8.3	19	5.5	20	2.8	8	2.0	17
40*F*	6.0	20	6.0	18	.0	20.5	∞	1.5
39*F*	5.5	21	4.4	21	.6	18	7.3	6

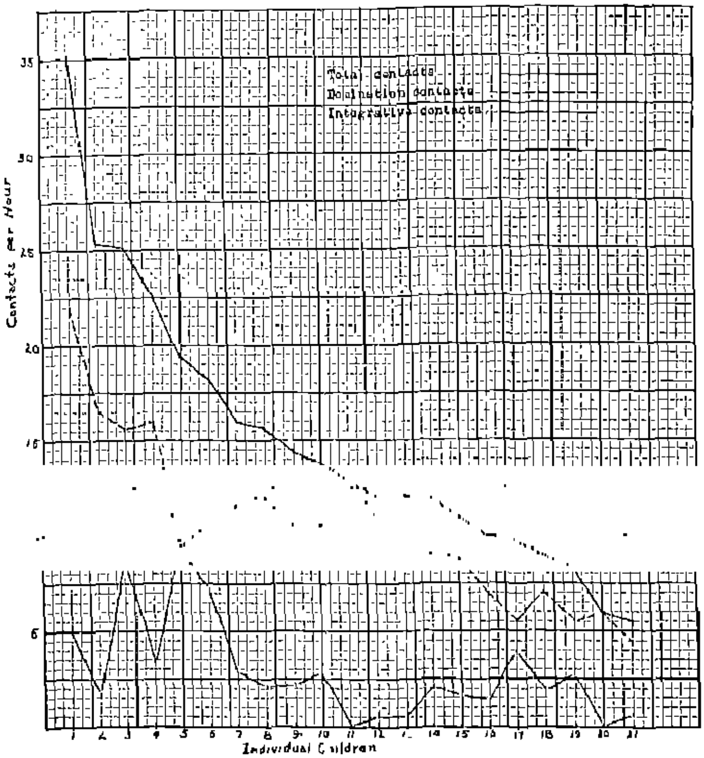


FIGURE 6

Mean Number of Contacts per Hour Which Teacher *B* Had with Individual Children Enrolled in the Afternoon Session (from Table 18)

The *D-I* ratios range from 0.7 to 5.5, with the median at 2.0. The child with the highest *D-I* ratio happens to be the median child in total number of contacts. In the morning group the child with the highest *D-I* ratio had the lowest frequency of total contacts.

d. Teacher B, afternoon session. Table 18 and Figure 6 show the mean number of contacts per hour which Teacher *B* had with the same individual children presented with Teacher *A* immediately above in Table 17 and Figure 5.

In ranges of contacts, Teacher *B* shows no considerable divergences from the data obtained for her with the morning group. The medians for the afternoon group are, however, considerably and consistently higher for total contacts and for domination and integration contacts.

The range of total contacts extends from 5.5 to 35.2, again with the first in rank order outstanding above the others. The median of 12.9 is 5.1 points higher than the median for Teacher *B*'s contacts with the younger children enrolled in the morning group.

Domination contacts range from 4.4 to 22.0, with the median of 10.5 standing in contrast with the median of 6.1 obtained for the morning group.

Integration contacts range from zero to 9.4, with the median at 2.1. Five children had less than one integrative contact per hour.

The *D-I* ratios extend from 1.0 to 18.3 for two children and to infinity for two more children who had zero integrative contacts. The median is 4.8. *The D-I ratios in the afternoon group are generally higher than those obtained with the same teacher in the morning group.*

e. Teacher C, morning session. Table 19 and Figure 7 give

TABLE 19

MEAN NUMBER OF CONTACTS PER HOUR WHICH TEACHER *C* HAD WITH INDIVIDUAL CHILDREN IN SCHOOL *Y*

Child	Total contacts Mean *N* per hr.	Rank	Domination Mean *N* per hr.	Rank	Integration Mean *N* per hr.	Rank	*D-I* ratio	*D-I* rank
48*M*	57.3	1	27.5	3	19.6	1	1.4	8.5
55*M*	55.3	2	35.8	1	15.3	2	2.3	6
53*F*	45.4	3	29.7	2	10.8	5	2.7	4
51F	41.9	4	26.8	4	11.0	4	2.4	5
49*F*	34.1	5	24.0	5	7.9	8	3.0	2
50*F*	33.0	6	19.1	8	9.7	6	2.0	7
52*M*	31.0	7	15.1	9	12.4	3	1.2	10
47*M*	30.1	8	22.3	6	5.9	11	3.8	1
54*F*	29.1	9	19.6	7	6.7	10	2.9	3
46*F*	23.0	10	10.1	10	9.3	7	1.1	11
45*F*	18.3	11	9.5	11	7.0	9	1.4	8.5

for Teacher *C* the mean number of contacts with individual children enrolled in School *Y*. A casual comparison of Table 19 and Figure 7 with Tables 15, 16, 17, and 18, and Figures 3, 4, 5, and 6, shows that the kindergartens in School *X* and School *Y* are different. It

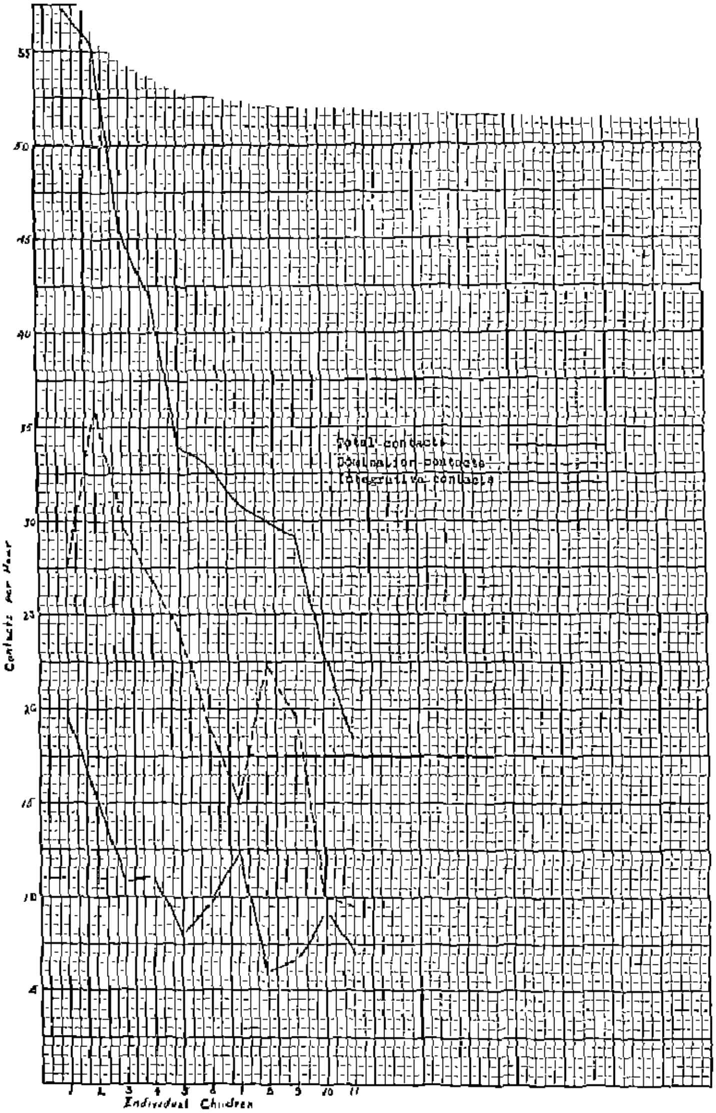


FIGURE 7
MEAN NUMBER OF CONTACTS PER HOUR WHICH TEACHER *C* HAD WITH INDIVIDUAL CHILDREN (FROM TABLE 19)

may be recalled that 23 children were enrolled in the morning session and 21 children in the afternoon session of School *X*. In School *Y* there were 11 children. The mean chronological ages of the three groups of children were 62.7 months for the morning group and 71.9 months for the afternoon group of School *X*, and 55.7 months for the 11 children in School *Y*. Three possible and uncontrolled variables may thus be operating to explain the large differences in teachers' techniques: the difference in mean age of the group, the difference in numbers in the group, and individual differences in the teachers themselves.

The striking contrasts between Teacher *C* and Teachers *A* and *B* suggest that an extension of the present research techniques to the measurement of the behavior of teachers in a larger number of schoolrooms selected for greater control of known variables would have important theoretical value for the psychology of human relations, and practical value for educators and mental hygienists.

The total number of teacher's contacts per hour in School *Y* ranges from 18.3 to 57.3. It will be noted that the two children highest in rank order had a mean number of contacts with Teacher *C* of approximately one a minute for a period of 440.5 minutes—a period of seven and one-third hours of observation. The child lowest in rank order with Teacher *C* was considerably above the median child with either Teacher *A* or Teacher *B*, whether in morning or afternoon sessions.

The domination contacts range from 9.5 to 35.8. As compared with Teacher *A*, Teacher *C* had frequencies of domination contacts per hour with four children each of whom exceeded the highest rank of Teacher *A* in the morning session, and with five children each exceeding the highest rank of Teacher *A* in the afternoon session. In dominative contacts, all children with Teacher *C* are each 50 per cent or more above the median for both morning and afternoon sessions with Teacher *A*.

The range of integrative contacts is from 5.9 to 19.6, with the median at 9.7. Compared with the children with Teacher *A*, the lowest in rank order among these 11 children is again above the median of either group in School *X*. The highest ranking child in School *Y* has over twice as many integrative contacts per hour as the highest ranking child in either group in School *X*.

The *D-I* ratios for the children with Teacher *C* are on the other hand in general comparable with the ratios of children with Teacher *A*; the range of 1.1 to 3.8 is slightly shorter at both ends than the ranges for both groups with Teacher *A*. The median of 2.3 is identical in one case and only slightly above the median of the other group for the children with Teacher *A*.

The *D-I* ratios of Teacher *C* are not similar to those of Teacher *B*, whose *D-I* ratios exceeded those of both Teacher *A* and Teacher *C*. It must be remembered that in School *X* the observers were recording *the contacts of that teacher who was playing the major rôle* with the children, and that Teacher *B* usually played a subsidiary rôle except during the music period when she took charge of the group. The relation of Teacher *C* to the children in School Y is more comparable to the rôle played by Teacher *A* who was the head teacher in School *X*.

In general the data show that the children with Teacher *C*, numbering approximately half those with Teacher *A*, have mean numbers of contacts per hour not quite double the frequencies of those with Teacher *A*. The practical importance of these observable and measurable differences must await subsequent research.

3. *Number of Teachers' Group Contacts per Hour According to Groups of Categories*

The contrasts and comparisons given above in Tables 15 to 19 inclusive represent the teacher's direct contacts with individual children. In addition to the frequencies listed in those tables there were many contacts which the teachers had with the children as a group, as for example, "*Boys and girls, listen to me*," or "*Who wants*

TABLE 20

FREQUENCIES PER HOUR OF TEACHERS' GROUP CONTACTS

				Categories				
School	Session	Teacher	Total obs. time (min.)	Total contacts 1-24	Domination 1-8	9-10	Integration 15-23	*D-I* ratio
X	A.M.	*A* & *B*	460.5	112.6	80.7	19.9	11.5	7.0
		A	362.0	93.0	65.5	16.9	10.0	6.6
		B	98.5	184.6	136.5	31.1	17.1	8.0
	P.M.	*A* & *B*	461.5	106.5	80.4	16.3	9.9	8.1
		A	352.5	84.1	60.4	14.6	9.0	6.7
		B	109.0	178.9	144.8	21.5	12.7	11.4
Y	A.M.	*C*	440.5	105.8	64.4	29.3	12.0	5.4

to be a robin?" The frequencies of these group contacts per hour are shown in Table 20.

It can be seen in Table 20 that Teacher *B* had approximately twice as many group contacts of all kinds, Categories 1-24, per hour of observation as did Teacher *A*, both for morning and for afternoon. When the data for total numbers of contacts for Teachers *A* and *B* are combined the new mean is not greatly in excess of the mean shown by Teacher *C*.

With group contacts as with individual contacts domination exceeds integration. The range of *D-I* ratios in Table 20 is from 5.4 to 11.4.

K. SEX COMPARISONS IN TEACHERS' CONTACTS WITH INDIVIDUAL CHILDREN IN SCHOOL *X*

The total number of contacts which Teachers *A* and *B* had with individual children in both morning and afternoon sessions were analyzed for sex differences. Table 21 shows for 19 boys and 25

TABLE 21

MEAN NUMBER OF TEACHERS' CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN ACCORDING TO SEX; DIFFERENCES BETWEEN THE MEANS, STANDARD ERRORS OF THE DIFFERENCES, AND CRITICAL RATIOS

Categories	Teacher	Boys $N = 19$		Girls $N = 25$	
		Mean	SE_m	Mean	SE_m
1-24	*A* and *B*	15.91	2.10	13.18	1.06
	A	16.61	2.24	13.79	1.23
	B	13.17	2.10	10.80	.97
1- 8	*A* and *B*	9.35	1.30	7.06	.44
	A	9.27	1.41	6.89	.50
	B	9.39	1.26	7.59	.64
15-23	*A* and *B*	4.39	.53	4.00	.44
	A	4.97	.65	4.57	.31
	B	2.40	.58	2.16	.44

Categories	Teacher	Higher mean	$Diff._m$	SE_{diff}	Crit. ratio	Chances in 100
1-24	*A* and *B*	Boys	2.73	2.35	1.16	87
	A	Boys	2.82	2.56	1.10	86
	B	Boys	2.37	1.83	1.30	90
1-8	*A* and *B*	Boys	2.29	1.37	1.67	96
	A	Boys	2.38	1.50	1.59	94
	B	Boys	1.80	1.42	1.27	89
15-23	*A* and *B*	Boys	.39	.62	.63	74
	A	Boys	.40	.52	.56	71
	B	Boys	.24	.43	.37	64

girls the means, standard errors of the means, obtained differences, standard errors of the differences, and the critical ratios for the total number of teachers' contacts, dominative contacts and integrative contacts. Without exception the boys have higher mean frequencies of contacts per hour regardless of the teacher or the nature of the contact.

For Teachers *A* and *B* combined and *A* and *B* separately there were 87, 86, and 90 chances in 100 that the respective obtained differences represented true differences greater than zero.

The critical ratios are even higher for domination contacts. For Teachers *A* and *B* combined and for Teachers *A* and *B* separately there are 96, 94, and 89 chances in 100 that the obtained higher mean frequencies of dominative contacts with boys represented respective true differences greater than zero.

For mean numbers of integrative contacts per hour the higher mean frequencies in favor of the boys are all small and the critical ratios in no case approach significance.

L. CORRELATION OF TEACHERS' DOMINATIVE CONTACTS WITH TEACHERS' INTEGRATIVE CONTACTS

In previous studies of the psychological interplay of children in an experimental play situation, domination contacts were found to have a zero correlation with integrative contacts (5, p. 344), (6, p. 397-402). It was of some interest to the experimenters to determine the relation of the domination contacts of the teacher to her integrative contacts. The children were arranged in rank orders of dominative and integrative contacts respectively for mean numbers of contacts per hour. Rank order correlations yielded the coefficients given in Table 22. All the coefficients are positive and within a fairly small range. The magnitude of the coefficients together with the consistency indicated by the range would suggest a tendency for the child who is high in integrative contacts to be high also in dominative contacts. This tendency toward a positive relationship between dominative and integrative contacts was demonstrated above in the general consistency of *D-I* ratios for individual contacts and for group contacts.

TABLE 22

RANK ORDER COEFFICIENTS OF CORRELATION BETWEEN TEACHERS' DOMINATIVE AND INTEGRATIVE CONTACTS BASED ON FREQUENCIES OF CONTACTS PER HOUR WITH INDIVIDUAL CHILDREN ARRANGED IN RANK ORDERS

School	*X*						*Y*
Session	A.M.			P.M.			A.M.
Teachers	*A* & *B*	*A*	*B*	*A* & *B*	*A*	*B*	*C*
r	.46	.54	.61	.59	.68	.48	.41

M. RELATION OF TEACHERS' CONTACTS TO MENTAL AGE AND CHRONOLOGICAL AGE OF THE CHILD

In order to determine whether or not teachers' contacts were related to mental age or to chronological age, correlations were computed between these respective ages and the mean numbers of teachers' individual contacts per hour. The coefficients are given in Table 23. The coefficients are small, ranging from .16 to .52 for mental age and from —.46 to .35 for chronological age, all with rather large probable errors.

TABLE 23

COEFFICIENTS OF CORRELATION BETWEEN THE MEAN NUMBERS PER HOUR OF TEACHERS' INDIVIDUAL CONTACTS WITH THE CHILD AND THE CHILD'S MENTAL AGE AND CHRONOLOGICAL AGE

Categories	Teacher	Numbers	Mental age	Chron. age
1-24	*A*, *B*, and *C*	48		—.45±.08
	A and *B*	37	.25±.10	—.15±.11
	A	37	.18±.11	.17±.11
	B	37	.47±.09	.17±.11
1-8	*A*, *B*, and *C*	48		—.44±.08
	A and *B*	37	.19±.11	—.15±.11
	A	37	.16±.11	—.12±.11
	B	37	.52±.08	.35±.10
15-23	*A*, *B*, and *C*	48		—.46±.08
	A and *B*	37	.21±.11	—.13±.11
	A	37	.19±.11	—.18±.11
	B	37	.22±.11	.14±.11

It can be noticed that all nine coefficients based on mental age, whether for dominative, integrative, or for total contacts, are positive. This by itself is consistent with the tendency reported above *for a positive relationship to exist between the teachers' dominative* and integrative contacts.

Of the 12 coefficients of correlation with chronological age eight of them are negative. The three coefficients that are over four times their probable errors are consistently negative and virtually identical. It may be said that when the contacts of Teachers *A*, *B*, and *C* are combined there is indicated a tendency for chronological age of the child to show an inverse relation to frequency of teachers' contacts, whether the contacts are dominative, integrative, or combined.

The consistency of the negative coefficients is also consistent with the tendency reported above for dominative contacts to be positively related to integrative contacts.

N. Summary and Conclusions

The purpose of the study reported in this section was to develop in terms of the concepts of dominative and integrative behavior measures of the observable contacts which teachers had with children. The methods were observational. Data were collected showing teachers' contacts with children in three different kindergarten schoolrooms. The subjects were three teachers and 55 children. The analysis of the data support the following statements.

1. Two independent observers showed a high degree of consistency in defining a teacher's contact and in their speed of recording. Five coefficients of consistency for total numbers of contacts per observation period during 73 observation periods of five minutes each were all .95 or above. During five hours and forty-two minutes of independent and simultaneous observation Observer *M* recorded 1,897 teachers' contacts and Observer *N* recorded 1,893.
2. Ten coefficients of consistency of two observers for teachers' contacts with individual children for total numbers of contacts of all kinds were .87 or above, six of the ten being .94 or above.
3. Ten coefficients of consistency of observers for teachers' dominative contacts with individual children (Categories 1-8) were .80 or above, six of the ten being .93 or above.
4. The coefficients of consistency of observers for teachers' integrative contacts with individual children (Categories 15-23) were based on lower frequencies and were low but consistently within a narrow range.
5. There was considerable evidence that in spite of the high degree of consistency of the observers, the observers were at times

unable to record the contacts at the speed with which they occurred. This would account in part for the lower coefficients of consistency for integrative contacts which by their nature must often be identified by their context, are therefore less specific and more difficult to record.

6. Analysis of the simultaneous records of two observers for consistency in recording individual categories showed coefficients of .85 to .96 for Category 1, which had by far the highest frequencies. Coefficients for other categories having sufficient numbers for statistical treatment showed 18 of 25 coefficients over .60. Pending further refinement of the individual categories the teachers' contacts are used in respective groups of domination and integration contacts.

7. The most rigorous method of analyzing all the data for consistency of two observers in which teachers' contacts were correlated child by child and category by category showed for 1,560 squares on record blanks for School *X* a coefficient of .78; and for 378 squares on record blanks for School *Y* a coefficient of .77. For this kind of data these coefficients are high.

8. The consistency of observers in recording the kind of data sought in this study is sufficiently high to be acceptable as measures of teachers' behavior.

9. No systematic tendencies for teachers' contacts to increase or decrease from one five-minute period to the next were found in those observation periods during which the activity could be satisfactorily defined.

10. A high degree of consistency in time samples for treatment of the data by category was found in periods totaling over 200 minutes of observation.

11. For a high degree of consistency in time samples for treatment of the data by groups of categories for individual children a total of from 300 to 400 minutes is sufficient.

12. Analysis of domination contacts of the head teacher, Teacher *A*, and the assistant teacher, Teacher *B*, showed differences that were statistically significant.

13. In comparing Teachers *A* and *B* as to frequency of contacts per child per hour rank order correlations of children for total numbers of contacts and for dominative and integrative contacts

yielded low coefficients with the highest coefficients for the entire observation period .48 for total contacts per hour.

14. In one group of 21 children the range of mean numbers of total teachers' contacts per hour was from 5.1 to 39.7, with the median child having 15.3 contacts per hour; in another kindergarten group of 11 children the range was from 18.3 to 57.3, with the median child at 33.0 contacts per hour.

15. In the group of 21 children Teacher *A* had a mean number of domination contacts per hour with individual children ranging from 3.4 to 23.2, with the median child at 6.5; in the group of 11 children Teacher *C* had mean numbers of individual domination contacts ranging from 9.5 to 35.8 per hour with the median child at 22.3.

16. Teacher *A* was found to have mean numbers of individual integrative contacts per hour ranging from .9 to 8.6 for 21 children, with the median child at 4.0; Teacher *C* had a range of mean numbers of individual integrative contacts of from 5.9 to 19.6 for 11 children, with the median child at 9.7.

17. For contacts with individual children the ratio of mean numbers of domination contacts per hour to mean numbers of integration contacts per hour (the *D-I* ratio) ranged from .7 to 5.5 for Teacher *A*, with the median at 2.0; for Teacher *B* the *D-I* ratios ranged from 1.0 to 18.3 and to infinity for two children who had no integrative contacts, with median child at 4.8; for Teacher *C* the *D-I* ratios ranged from 1.1 to 3.8, with the median at 2.3.

18. The mean numbers of teachers' contacts per hour differentiate teachers and differentiate children.

19. The *D-I* ratios differentiate both children and teachers.

20. The mean numbers of "group" contacts per hour show a frequency for Teacher *B* twice that of Teacher *A* for total number of contacts and for domination contacts.

21. Teacher *C*, with 11 children had a considerably higher mean number per hour of group contacts of all kinds than did either Teacher *A* or Teacher *B* with 23 children and 21 children in morning and afternoon sessions respectively.

22. The *D-I* ratios for teachers' mean number of group contacts per hour exceeded for Teachers *A* and *C* but not for Teacher *B* the *D-I* ratios for mean numbers of contacts per hour with individual children. The range of *D-I* ratios for group contacts per

hour by individual teachers began with the *D-I* ratio of 5.4 for Teacher *C*, and extended through 6.6 and 6.7 for Teacher *A* respectively for morning and afternoon sessions, to 8.0 and 11.4 for Teacher *B* respectively for morning and afternoon sessions.

23. From Teachers *A* and *B* the boys in all cases regardless of the nature of the contacts received higher mean frequencies of teachers' contacts than did the girls. None of the differences was statistically significant, but in the case of dominative contacts and total contacts three of six critical ratios approached significance.

24. Rank correlations between teachers' dominative and integrative contacts based on frequencies of individual contacts per hour, with children arranged in rank order, yielded seven coefficients ranging from .41 to .68.

25. The coefficients of correlation between teachers' dominative and integrative contacts together with the consistency of *D-I* ratios is evidence of a tendency toward a positive relationship between dominative and integrative techniques for the teachers in this study. This positive relationship is further supported by the consistency in signs of respective coefficients of correlation when both dominative and integrative contacts are correlated respectively with mental age and chronological age.

26. Of nine coefficients of correlation between mean numbers of teachers' individual contacts per hour and mental age, only two are over four times their probable errors: Teacher *B*'s dominative contacts have a coefficient of .52±.08, and Teacher *B*'s total contacts have a coefficient of .47±.09 with mental age. All other coefficients of correlation of dominative contacts, integrative contacts, or of total contacts with mental age are positive but small, ranging from .16 to .25.

27. Chronological age of the child shows consistently negative and virtually identical coefficients of correlation with mean number of total contacts per hour (—.45±.08), mean number of dominative contacts per hour (—.44±.08), and mean number of integrative contacts per hour (—.46±.08); all three coefficients being based on the combined data for Teachers *A*, *B* and *C*.

III. DOMINATION AND INTEGRATION IN THE SOCIAL BEHAVIOR OF KINDERGARTEN CHILDREN IN AN EXPERIMENTAL PLAY SITUATION

A. Aims

The purposes of this part of the investigation may be stated as follows:

1. To determine whether experimental techniques previously developed for studying domination and integration in the social interplay of preschool children would be applicable to the study of kindergarten children.
2. To determine which findings previously reported for children of preschool age would be supported by a similar study of kindergarten children, and which of the previous findings would not be consistent with the behavior of kindergarten children.
3. To determine the relationship between domination and integration in the teachers' contacts with kindergarten children in the schoolroom and domination and integration scores in child-child relationships in this experimental situation.

B. Methods and Procedure

An attempt was made to reproduce as exactly as possible the experimental situation and procedure adopted for the previous study with preschool children (5, pp. 339-340), (6, pp. 349-350). Children were taken in pairs to the testing room where they were allowed to play for five minutes. In the room was a sand box on a low table. In the sand box were a toy sand pail, shovel, sieve, two automobiles, and three rubber toy animals always arranged in the same position when the children arrived. As he led the children into the room, the experimenter said:

> Here are some toys for you to play with until I come back and get you. We will keep them in the sand box all the time, but you may play with anything you want to.

In two of the kindergarten groups each child with a few exceptions was paired at random with five other children of his group. With a few exceptions in the cases of children who had been absent a great deal no child was paired more than twice on the same day.

In general, the children expressed considerable interest in coming to the experimental room and many were frankly enthusiastic.

C. THE PORTABLE TESTING BOOTH

For the paired experiments a portable experimental booth was constructed. It was made of three-ply wood with two-by-two supports at the corners. The booth was six feet by eight with a hinged door at one end. On one side were two one-way observation screens and an observation booth four by eight feet in size. The "play room," as it was called with the children, had no ceiling over it, but the attached observation booth had a ceiling that was light proof. The side panels were held together by turnbuckles in the corners at the top and bottom. The portable experimental booth could be dismantled in one school, moved to another and set up again complete with observation booth attached in less than two hours. In School *X* the booth was set up for a time in a corner of the gymnasium which was not in use during the experiments. One week when the gymnasium was to be used several times the booth was taken down and moved to an enclosed corridor. In School *Y* the portable experimental booth was set up in the "music room" which had no fixed seats in one-half of the room. When placed directly under a light in the school building, accessory lighting for the one-way observation screens was unnecessary.

For practically all the pairings one of the experimenters served as a "messenger" who called for the children and returned them at the end of five minutes. This made it possible for the observers to be stationed in the observation booth before the children arrived. An observer would arrange the toys in the sand box while the messenger was getting the next pair of children. The messenger stationed herself at some distance from the booth and "worked" while the children were in the experimental booth. The children could hear her walk away and hear her return for them. Of all the children who supplied data for this experiment only two gave evidence of having discovered the presence of persons in the observation booth. Many suspected that the screens were windows, but the screens were referred to as windows, blackboards, and pictures on the wall.

D. SUBJECTS

The children participating in this experiment were 49 of the 55

TABLE 24
NUMBERS OF CHILDREN USED IN THE PAIRED COMPARISONS

School	Session	Boys	Girls	Total	Total pairings
X	A.M.	8	11	19	30
	P.M.	8	11	19	96
Y	A.M.	5	6	11	50
Totals		21	28	49	176

kindergarten children reported in Table 1. Table 24 gives the distribution of boys and girls according to kindergarten groups together with the total number of pairings in each group. In the morning group of School *X* most children had two pairings. In the afternoon group all children with few exceptions were paired with five others. In School *Y* also most of the children were paired five times.

E. METHODS OF SCORING

The observation blank shown in Figure 8 is based on the blank which was developed for the previous study (6, pp. 350-351). Domination scores in a pairing included all tallies in Items 1-8 inclusive. Integration scores included all tallies for Items 9-12 inclusive plus the number of the companion's tallies for Item 11.

F. DEFINITIONS

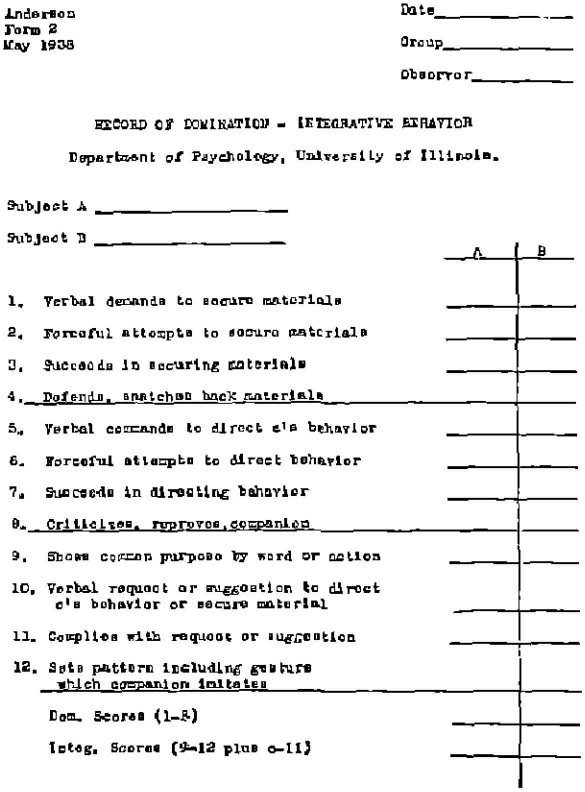
The definitions of categories of domination and of integrative behavior in children's responses to each other were taken exactly from the previous study and are reproduced herewith:

1. *Verbal Demands to Secure Materials*

Demands toy or play materials (including the right to use the sand in the box or in a certain part of the box).

2. *Forceful Attempts to Secure Materials*

Attempts to get toy or play materials (including certain position at or certain portion of the sand box) from companion. Includes attempts to take material even when companion has laid it down close to him, but is obviously retaining it in his immediate possession. It includes, also, reaching for material even when the hands do not come into actual contact with it before it is snatched away. (Forceful accompaniments to attempts to get materials or to enforce other demands, as hitting and kicking, are not counted again as separate items of be-



Anderson
Form 2
May 1938

Date____________
Group____________
Observer____________

RECORD OF DOMINATION - INTEGRATIVE BEHAVIOR

Department of Psychology, University of Illinois.

Subject A ____________

Subject B ____________

	A	B
1. Verbal demands to secure materials		
2. Forceful attempts to secure materials		
3. Succeeds in securing materials		
4. Defends, snatches back materials		
5. Verbal commands to direct c's behavior		
6. Forceful attempts to direct behavior		
7. Succeeds in directing behavior		
8. Criticizes, reproves, companion		
9. Shows common purpose by word or action		
10. Verbal request or suggestion to direct c's behavior or secure material		
11. Complies with request or suggestion		
12. Sets pattern including gesture which companion imitates		
Dom. Scores (1-8)		
Integ. Scores (9-12 plus c-11)		

FIGURE 8

havior.) (When 1 and 2 occur simultaneously, only Item 2 is counted.)

3. *Succeeds in Securing Materials from Companion's Possession by Demands or by Force*

4. *Defends, Snatches Back Materials Taken from His Possession*

Holds on to toy he has and resists attempts to take it; snatches it back from companion who has taken it from him. Item 4 is counted even when the companion ultimately succeeds in getting the toy, but is counted only once for each attempt to take material from a child.

Defense is construed here to involve a command, threat, or the use of force.

5. *Verbal Commands to Direct Companion's Behavior*

Includes resistance to companion's attempts to participate in child's activity.

One child forbids another.

6. *Forceful Attempts to Direct Behavior*

Forceful attempts usually include contact with companion's person or materials. Child *A* saying, "Look what I made" is not credited with forceful attempt but with Item 10, verbal suggestion to direct behavior. However, if the companion pays no attention and the child holds the object in his cmpanion's line of vision, this is regarded as a forceful attempt to direct behavior.

Throwing sand or other material at companion; use of force of any kind about the person of companion. (When 5 and 6 occur simultaneously, only Item 6 is counted.)

7. *Succeeds in Directing Behavior by Commands or Force*

(Success in Items 5 or 6.)

8. *Criticizes, Reproves, or Places Companion in a Disadvantageous Comparison*

Includes "Bum," "Dumbbell," and other uncomplimentary names.

This behavior implies an attack against the status of a companion and includes, therefore, any implication of blame. "Criticizes" implies negative criticism and does not include criticism that is constructive or that attempts to evaluate a product or activity independently of the companion associated at the moment with it.

Behavior in which child places companion in disadvantageous comparison with self is not construed as evidence that child is seeking harmony or common purpose, e.g., "I have a bigger house than your's," "I saw a bigger elephant than you did." Such remarks when obviously stated not to compare houses or elephants so much as to compare Child *A* with Child *B* are construed as dominating and checked under Item 8.

9. *Shows Common Purpose by Word or Action*

Includes attempts to coöperate; attempts to share or to participate in companion's activity, e.g., "Now *I'll* do this": also acquiescence to such attempts on the part of companion. Non-verbal behavior in the nature of taking turns. Such statements as, "Let's cover him up," when the companion is already engaged in the suggested activity; instances when the child may help his companion with something in order to show how to do it, although there is no verbal accompaniment to indicate that this is the reason. Without overt evidence of domination the use of *"We"* presumes integrative behavior.

Gratuitous behavior; volunteering services; giving materials.

"This is the way I do it" (am going to do it).

Acquiescence in or overt recognition of, or response to, coöperative behavior, volunteered services or donated materials.

10. *Verbal Request or Suggestion to Direct Companion's Behavior or to Secure Materials*

Child suggests activity for his companion, either individual or coöperative. This includes any statement for companion's activity beginning with *"Let's"* or suggestion in form of question as *"Shall we."* (Note under 9 that *"Let's"* indicates common purpose when companion is himself already engaged in the activity; the change in behavior is then in the child himself. In 10 *"Let's"* implies that the suggested change in behavior is primarily in the companion, the child himself already having the idea.)

Includes apportionment of toys to companion with or without a verbal accompaniment (such statements as, "Do you want me to pour some sand into your truck?").

Such vocalization as *"E-e-e-e"* when obviously directing companion to stop doing something is also included.

Includes all questions: e.g., "When will you be through with the shovel?" "What are you making?" "Did you hear that noise?" "What did you say?" "Huh?" "What?"

Includes all remarks calling child's attention to something: e.g., "Look at this," "Watch me." (These are construed suggestions unless accompanied by evidence of force or threat.)

Child asks for or suggests that he be given toy or play materials (including the right to use the sand in the box or in a certain part of the box); includes any statement or desire for something in the possession of companion and modification of such statements as: "Can I have it, if I give it back in just a moment?"

Unrelated conversation, not questions or answers: e.g., "I went to the circus," "I've got a blue dress."

11. *Complies with Request or Suggestion*

Includes answers, gestures, or overt responses to any question; includes repetition of an original question when companion asked, e.g., "What did you say?"

Repetition of a remark when companion asks, "What did you say?" Repetition of an act when companion says, "What did you do?" or "Do that again."

12. *Sr's Pattern Including Gesture which Companion Imitates*

This is counted only when there is repetition of the speech or other behavior of the companion, the pattern being repeated in sufficient entirety to be plainly recognizable. Only one repetition of the same pattern is counted. (It does not include joining companion with a song or chanted phrase, which would be checked as 9.)

Example. The sand table jiggled. One child stooped down to look at underside of table; companion did same. *Example.* Child goes to window and looks out; companion does same (6, pp. 351-354).

G. CONSISTENCY OF OBSERVATIONS BY TWO OBSERVERS[3]

The coefficients of consistency of simultaneous and independent recording by two observers are given in Table 25. The coefficients are .89±.02 for domination scores and .90±.02 for integration scores.

TABLE 25

COEFFICIENTS OF CONSISTENCY OF SIMULTANEOUS RECORDING BY OBSERVERS *O* AND *P*

	Pairings	r	PE_r
Integration	70	.90	.02
Domination	78	.89	.02

H. SEX DIFFERENCES

1. *Résumé of Previous Findings*

In the previous study of children of preschool age three groups of children were used: one group of Iowa City children superior in intelligence and coming from homes of superior socio-economic status, and two groups of children in an orphanage, one of which (preschool group) attended a nursery school, the other (control group) living in the orphanage but not attending nursery school. When data from all three groups were combined to show sex comparisons of mean domination scores and mean integration scores it was found that girls were more dominating than boys, the critical ratio between the obtained difference and the standard error of the difference being 4.60. It was found also that girls were less integrative than boys, there being 98 chances in 100 that the obtained difference represented a true difference greater than zero.

---

[3] In order to determine the consistency of his recording of domination and integration behavior the writer observed simultaneously with Gladys Lowe Anderson.

Other findings of sex differences in the orphanage children alone may be summarized as follows:

a. The nursery school and control groups of orphanage children showed a consistency of scores in their sex comparisons; they differed only in degree.

b. Girls were more dominating than boys.

c. When sexes were cross-paired, girls decreased in dominative behavior and boys increased in dominative behavior, the difference in the former case being significant; in the latter case there were 92 chances in 100 that the obtained difference represented a true difference greater than zero.

d. Girls and boys were not significantly different in integrative behavior when paired with their own sex.

e. When the sexes were cross-paired, both boys and girls showed decreases in integrative behavior.

f. Boys were slightly more integrative in the cross-pairings of sexes than were girls, though the difference was not significant.

g. In cross-pairings, girls showed lower scores than boys in both dominative and integrative behavior. This was true for both nursery school and control groups of children.

h. As regarded total interactivity (the sum of domination and integration scores), the data presented indicated that among these small orphanage children sex differences were already established (6, pp. 376-388).

2. *Findings of the Present Study*

Table 26*a* gives for the kindergarten children in the present study the mean domination and mean integration scores respectively, together with standard errors of the means and Domination-Integration ratios, for boys when paired with boys, boys when paired with girls, girls when paired with boys, and girls when paired with girls.

TABLE 26*a*

SEX COMPARISONS IN DOMINATION AND INTEGRATION SCORES

Means, standard errors of the means, and Domination-Integration ratios for boys when paired with boys, boys when paired with girls, girls when paired with boys, and for girls when paired with girls.

		Domination		Integration		
Score for	Number pairings	Mean	SE_{mean}	Mean	SE_{mean}	*D-I* ratio
School X, A.M.						
Boys with boys	6	2.83	1.58	47.67	7.69	.06
Boys with girls	6	3.00	1.60	31.00	5.29	.10
Girls with boys	6	1.17	.55	33.83	8.67	.04
Girls with girls	12	1.67	1.30	32.50	4.09	.05
School X, P.M.						
Boys with boys	16	4.50	1.35	33.88	3.19	.13
Boys with girls	25	2.88	.73	35.20	2.10	.08
Girls with boys	25	3.28	.75	35.40	1.65	.09
Girls with girls	30	1.33	.41	38.63	1.78	.04
School X, A.M. + P.M.						
Boys with boys	22	4.05	1.09	37.64	3.38	.11
Boys with girls	31	2.90	.68	34.39	2.00	.08
Girls with boys	31	2.87	.67	35.10	1.82	.08
Girls with girls	42	1.43	.33	36.88	1.77	.04
School Y, A.M.						
Boys with boys	8	1.63	.92	29.63	2.71	.06
Boys with girls	15	1.20	.32	30.33	2.42	.04
Girls with boys	15	1.13	.34	33.07	2.87	.03
Girls with girls	10	1.50	.60	34.20	2.38	.04

Table 26*b* shows for different comparisons of mean scores the obtained differences between mean scores, standard errors of the differences, critical ratios, and the chances in 100 that the obtained differences represent true differences greater than zero.

a. Domination. Figure 9 shows graphically the mean domination scores given in Table 26*a*, together with a reproduction to scale of the mean domination scores of the orphanage nursery school and control groups taken from Table 21 and Figure 8 of that study (6, pp. 380, 379).

(1). *Résumé of the domination scores of orphanage children. In the previous study of orphanage children it was found that when* a more dominating group was cross-paired with a less dominating group there was a dynamic relation which was sufficient to indicate a tendency toward over-compensation. This dynamic relation to-

TABLE 26*b*

Sex Comparisons from Table 26*a*

Differences between the means, standard errors of the differences, critical ratios, and chances in 100 for domination and integration respectively.

Group	Higher mean	Difference	Standard error of difference	Critical ratio	Chances in 100
		Domination			
School X, A.M.					
**BB* with *GG*	*BB*	1.16	2.05	.57	72
BG with *GB*	*BG*	1.83	1.69	1.08	86
BB with *BG*	*BG*	.17	2.25	.08	53
GB with *GG*	*GG*	.50	1.41	.35	64
BB with *GB*	*BB*	1.66	1.67	.99	84
BG with *GG*	*BG*	1.33	2.06	.65	74
School X, P.M.					
BB with *GG*	*BB*	3.17	1.41	2.25	99
BG with *GB*	*GB*	.40	1.10	.36	64
BB with *BG*	*BB*	1.62	1.53	1.06	86
GB with *GG*	*GB*	1.95	.85	2.29	99
BB with *GB*	*BB*	1.22	1.54	.79	79
BG with *GG*	*BG*	1.55	.84	1.85	97
School X, A.M. + P.M.					
BB with *GG*	*BB*	2.62	1.14	2.30	99
BG with *GB*	*BG*	.03	.96	.03	52
BB with *BG*	*BB*	1.15	1.28	.90	82
GB with *GG*	*GB*	1.44	.75	1.92	97
BB with *GB*	*BB*	1.18	1.28	.92	82
BG with *GG*	*BG*	1.47	.76	1.93	97
School Y, A.M.					
BB with *GG*	*BB*	.13	1.10	.12	55
BG with *GB*	*BG*	.07	.46	.15	56
BB with *BG*	*BB*	.43	.97	.44	67
GB with *GG*	*GG*	.37	.69	.54	71
BB with *GB*	*BB*	.50	.98	.51	69
BG with *GG*	*GG*	.30	.68	.44	67
School X, A.M. compared with P.M.					
BB with *BB*	P.M.	1.67	2.08	.80	79
BG with *BG*	A.M.	.12	1.76	.07	53
GB with *GB*	P.M.	2.11	.93	2.27	99
GG with *GG*	A.M.	.34	1.36	.25	60

*The symbols "*BB* with *GG*" indicates that the mean score for boys when paired with boys is being compared with the mean score for girls when paired with girls. "*BG* with *GB*" indicates that the mean score for boys when paired with girls is being compared with the mean score for girls when paired with boys.

TABLE 26*b* (*continued*)

Group	Higher mean	Difference	Standard error of difference	Critical ratio	Chances in 100
		Integration			
School X, A.M.					
BB with *GG*	*BB*	14.97	8.71	1.72	96
BG with *GB*	*GB*	2.83	10.16	.28	66
BB with *BG*	*BB*	16.67	9.33	1.79	96
GB with *GG*	*GB*	1.33	9.59	.14	56
BB with *GB*	*BB*	13.84	11.59	1.19	88
BG with *GG*	*GG*	1.50	6.69	.22	59
School X, P.M.					
BB with *GG*	*GG*	4.75	3.65	1.30	90
BG with *GB*	*GB*	.20	2.39	.08	53
BB with *BG*	*BG*	1.32	3.82	.35	64
GB with *GG*	*GG*	3.23	1.58	2.04	98
BB with *GB*	*GB*	1.52	3.59	.42	66
BG with *GG*	*GG*	3.43	2.40	1.43	92
School X, A.M. + P.M.					
BB with *GG*	*BB*	.76	3.80	.20	58
BG with *GB*	*GB*	.71	2.70	.26	60
BB with *BG*	*BB*	3.25	3.91	.83	79
GB with *GG*	*GG*	1.78	2.54	.70	76
BB with *GB*	*BB*	2.54	3.82	.66	74
BG with *GG*	*GG*	2.49	2.67	.93	83
School Y, A.M.					
BB with *GG*	*GG*	4.57	3.61	1.27	89
BG with *GB*	*GB*	2.74	3.75	.73	77
BB with *BG*	*BG*	.70	3.63	.22	59
GB with *GG*	*GG*	1.13	3.73	.30	62
BB with *GB*	*GB*	3.44	3.95	.87	81
BG with *GG*	*GG*	3.87	3.39	1.14	87
School X, A.M. compared with P.M.					
BB with *BB*	A.M.	13.79	8.32	1.66	5
BG with *BG*	P.M.	4.20	5.68	.74	77
GB with *GB*	P.M.	1.57	8.82	.18	56
GG with *GG*	P.M.	6.13	4.46	1.37	91

gether with the over-compensation was shown both in the cross-pairings of the more dominating orphanage nursery school children with the less dominating control group children and in the cross-pairings of the more dominating girls with the less dominating boys. In the cross-pairings of the sexes the orphanage control group showed the same tendencies as the orphanage nursery school children, though

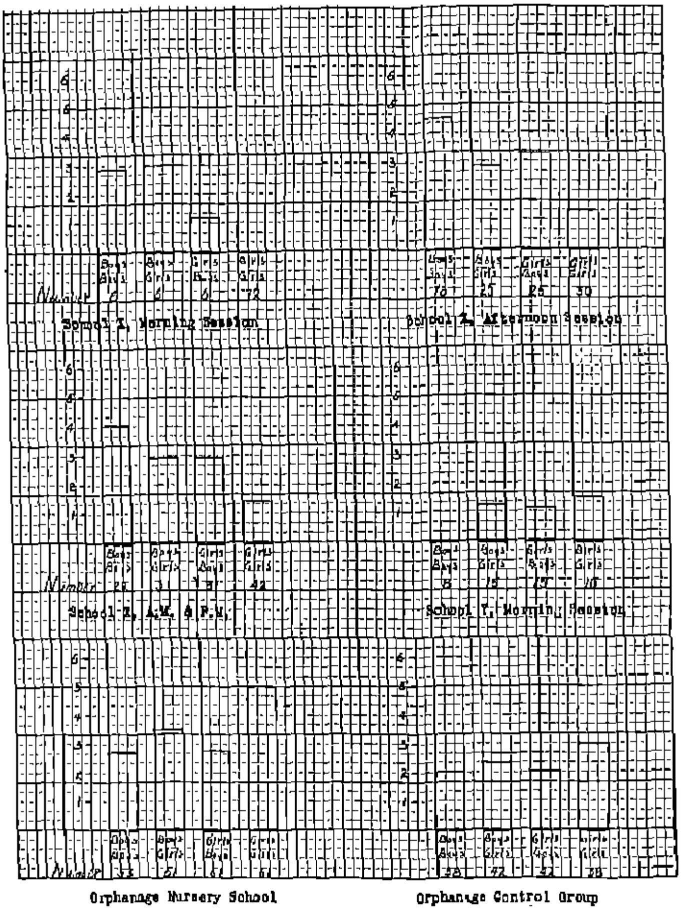


FIGURE 9

Mean Domination Scores for Boys When Paired with Boys, Boys When Paired with Girls, Girls When Paired with Boys, and Girls When Paired with Girls

Charts for orphanage children adapted from Anderson (6, pp. 379-380.)

the changes did not represent significant differences. By way of explanation it was said in the previous report that: "When the little sheep is paired with the big, bad wolf, the little sheep becomes not only less sheepish than the sheep but more wolfish than the wolf. And the wolf becomes more sheepish than the sheep" (6, p. 368).

(2). *School X, morning session.* In the morning session of School *X* the numbers of pairings were so small as not to offer sufficient data for statistical treatment. None of the sex differences in this small group of numbers even approximated significance.

(3). *School X, afternoon session.* In the afternoon the numbers of pairings were considerably larger as it was possible to complete the pairings of practically all the children in attendance with five others. In contrast, however, with the data of the previous study with orphanage children the kindergarten boys when paired with boys were more dominating than the kindergarten girls when paired with girls. There are 98 chances in 100 that the obtained difference represents a true difference greater than zero.

In spite of the reversal of sex supremacy established in the previous study for domination scores in own pairings, it can be seen by inspection in Figure 9 that among these kindergarten children there is again present in the cross-pairings both the dynamic relationship and the tendency toward over-compensation.

When the girls were cross-paired with boys their mean domination scores increased and the mean domination scores for the boys decreased as compared with the respective mean scores when paired with their own sex. The increase for the girls approaches significance, there being 94 chances in 100 that the obtained difference represents a significant difference. The decrease in the boys' mean domination scores does not approach significance, nor do any of the other sex comparisons in this group.

(4). *School X, morning and afternoon groups combined.* When the data for the fewer pairings in the morning session were combined with the data for the pairings in the afternoon session, the dynamic relationship of dominative behavior to dominative behavior was still present although there was no tendency toward over-compensation in the cross-pairings.

A glance at the domination scores represented in Figure 9 shows that the lowest mean domination scores were obtained by girls when paired with girls. The differences between these scores and the domination scores represented by the three other columns approaches statistical significance in each case. As compared with mean domination scores of boys when paired with boys there are 99 chances in

100 that the obtained difference represents a true difference greater than zero. As compared with the scores for girls when paired with boys there are 97 chances in 100 that the obtained difference is statistically significant.

(5.) *School Y, morning session.* In School Y there were no sex differences in mean domination scores that have an even chance of being significant. It may be pointed out, however, that the numbers were small.

(6). *Sex comparison of mean domination scores of children enrolled respectively in morning and afternoon sessions, School X.* The children enrolled in the afternoon session were older than the children enrolled in the morning group. When boys were paired with boys or when girls were paired with girls the differences between the morning and afternoon groups of children are not significant. Nor was there a difference when boys were paired with girls. However, when girls were paired with boys, there was a greater mean domination score for the girls enrolled in the afternoon group that has 98 chances in 100 of representing a true difference greater than zero.

b. Integration. Figure 10 shows graphically the mean integration scores given in Table 26*a*, together with a reproduction to scale of the mean integration scores of the orphanage nursery school and control groups taken from Table 21 and Figure 9 of that study (6, pp. 380, 384). In comparing domination scores with integration scores it should be noted that Figures 9 and 10 are not drawn to the same vertical scale. One unit of measurement in Figure 9 is equal to six units on Figure 10.

(1). *Résumé of the study of the orphanage children.* In the study of orphanage children the nursery school group was found to be less integrative than the control group in the comparison of homogeneous or own-group pairings. In the cross-pairings between nursery and control group children the same dynamic tendencies toward over-compensation were found as was revealed in dominative behavior: the low integrative group increased in integration scores and the high integrative group decreased so that the relationship of low to high was reversed.

In the cross-pairing of sexes, however, the dynamic tendency toward over-compensation found in mean domination scores was absent in integrative behavior. In both nursery school and control groups

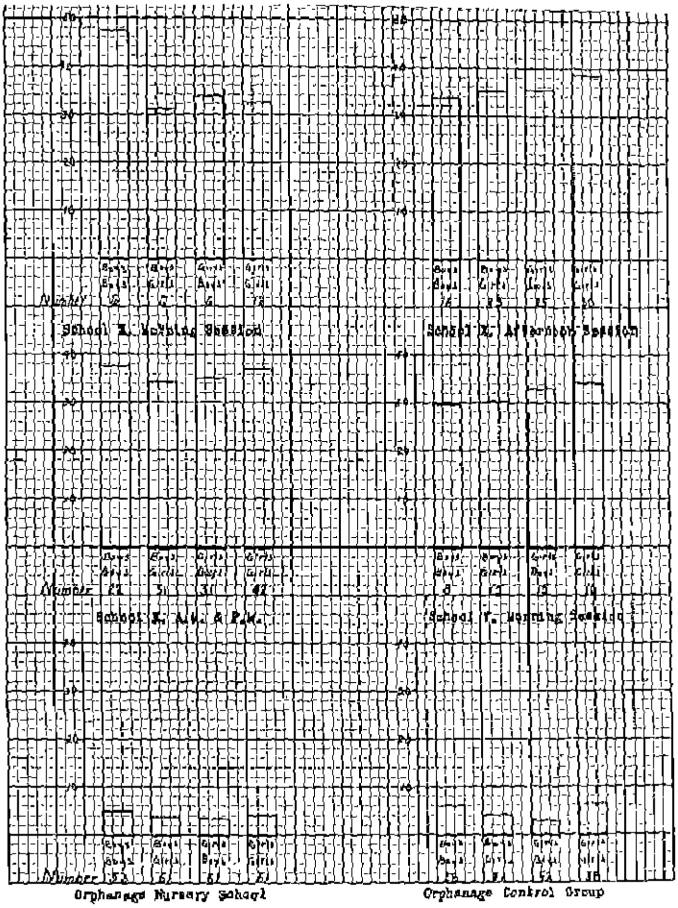


FIGURE 10

Mean Integration Scores for Boys When Paired with Boys, Boys When *Paired with Girls*, *Girls* When Paired with Boys, and Girls When Paired with Girls

Charts for orphanage children adapted from Anderson (6, pp. 380, 384).

the boys when paired with boys were not significantly different from girls when paired with girls. In the cross-pairings of the sexes the mean integration scores of both boys and girls dropped for both the nursery school and control group orphanage children. In the control group the drop in boys' mean score had 96 chances in 100 of representing a significant difference; the drop in girls' mean integration score in the cross-pairings had 99 chances in 100 of representing a true difference greater than zero. A justifiable inference is that these decreases in integrative behavior represent an inhibition or shyness in the cross-pairings not present in the own-sex pairings.

(2). *School X, morning session.* Although the numbers of pairings in this group were small, the boys when paired with boys had a higher mean integration score than did the girls when paired with girls, with 91 chances in 100 that the obtained difference represents a significant difference. The graph for this group shows the apparent tendency toward dynamic over-compensation in the cross-pairings. The increase in the girls' mean score, however, is not significant although the decrease in the boys' mean integration score has 93 chances in 100 of representing a difference greater than zero. It must be pointed out again, however, that the number of pairings in this group was small.

(3). *School X, afternoon session.* In this group the boys when paired with boys were lower in integration scores, though not significantly so, than were the girls when paired with girls. In the cross-pairings the boys gained in integration scores though the gain was not significant; the girls, however, showed a decrease that has 96 chances in 100 of representing a true difference greater than zero.

(4). *School X, morning and afternoon sessions combined.* When the integration scores for morning and afternoon groups were combined for greater numbers, the mean scores for own-sex pairings of boys and girls are not significantly different. In the cross-pairings the mean scores for both sexes dropped. To this extent the data are similar to the data obtained for both orphanage groups. The decreases in mean integration scores are, however, not significant.

(5). *School Y, morning session.* In School Y none of the sex comparisons of mean integration scores showed differences that even approach significance. The numbers were, however, small.

(6). *Sex comparisons of mean integration scores of children enrolled respectively in morning and afternoon sessions.* When the children enrolled in the morning session were compared with the children in the afternoon session no significant differences were found in mean integration scores. The younger boys enrolled in the morning session had a mean score considerably above the mean score for the afternoon group when boys were paired with boys. The standard error of the mean for the morning group was large, however. The difference has 90 chances in 100 of representing a true difference greater than zero. When girls were paired with girls, the reverse relation was found between morning and afternoon groups. The afternoon group of girls had a higher mean integration score. There are, however, only 83 chances in 100 that this obtained difference represents a true difference greater than zero.

I. DOMINATION-INTEGRATION RATIOS

D-I ratios are obtained by dividing domination scores by integration scores. In the measurement of teachers' contacts with children it was found that with very few exceptions the teachers used dominating techniques much more frequently than they used integrative techniques. Table 26*a* shows that in the contacts which these kindergarten children had with each other the reverse ratio obtained. With the exception of the behavior of the older boys in the afternoon session of School *X* there are no mean scores in which domination is even one-tenth the mean integration score.

The *D-I* ratios of these kindergarten children were different also from those shown in the previous study of children in the orphanage. In that study the *D-I* ratios were not presented as such, but they were easily obtained (6, p. 380) and for comparison with the present study are given in Table 27. For the orphanage nursery school children the *D-I* ratios ranged from .54 to 1.19, and for the

TABLE 27

MEAN DOMINATION SCORES, MEAN INTEGRATION SCORES, AND DOMINATION-INTEGRATION RATIOS OF ORPHANAGE CHILDREN FOR BOYS WHEN PAIRED WITH BOYS, BOYS WHEN PAIRED WITH GIRLS, GIRLS WHEN PAIRED WITH BOYS, AND GIRLS WHEN PAIRED WITH GIRLS

(Adapted from data given in 6, p. 380.)

	Orphange nursery school			Orphanage control		
	Mean dom.	Mean integ.	*D-I* ratio	Mean dom.	Mean integ.	*D-I* ratio
Boys with boys	2.70	5.02	.54	2.02	6.00	.34
Boys with girls	3.46	3.90	.89	2.33	4.02	.58
Girls with boys	2.79	3.57	.78	2.02	2.88	.70
Girls with girls	4.97	4.18	1.19	2.97	6.03	.49

orphanage control group the *D-I* ratios ranged from .34 to .70. These ratios are without exception higher than those obtained with the kindergarten children in the present study. It should be mentioned that almost without exception they are in turn not only lower but considerably lower than the *D-I* ratios for teachers' contacts with these same kindergarten children.

J. CONSISTENCY OF DOMINATION AND INTEGRATION SCORES BETWEEN FIRST AND FIFTH PAIRINGS

In seven cases of children who received their fourth and fifth pairings on the same day it was impossible to identify the pairing which came last. Both pairings were therefore averaged to obtain the mean for comparison with the domination and integration scores on the first pairing. Table 28 gives for the first and fifth pairings the mean domination scores, the respective standard errors of the means, the obtained differences, standard errors of the differences, and critical ratios. The obtained difference is statistically significant for mean domination scores, and for mean integration scores there are 96 chances in 100 that the obtained difference represents a true difference greater than zero. By way of interpretation it may be stated that it was an unusual experience for the children to be permitted to leave the kindergarten rooms to play by themselves. During the first pairings some of the children talked only in whispers. For most children who showed a tendency to talk in whispers the whispering disappeared during the first or second pairing. It may be noted that for similar experiments in the public

TABLE 28

COMPARISON OF MEAN DOMINATION AND MEAN INTEGRATION SCORES OF FIRST PAIRINGS RESPECTIVELY WITH FIFTH PAIRINGS

	Domination		*Integration*	
	Mean	SE_{mean}	Mean	SE_{mean}
1st pairing	1.11	.38	31.56	1.81
5th pairing	3.87	.91	36.61	1.71

	Obtained difference	$SE_{diff.}$	Critical *ratio*	Chances in 100
Domination	2.76	.99	2.79	100
Integration	5.05	2.49	2.03	98

schools where testing is an unusual procedure some preliminary adaptation to the portable testing room should be provided.

K. Correlations of Child's Domination and Integration Scores with Teacher's Contacts and Other Factors

Table 29 gives the coefficients of correlation between the child's domination scores, the child's integration scores and the child's total interactivity scores on the one hand and the teachers' contacts and the child's mental and chronological ages on the other.

TABLE 29

Coefficients of Correlation between Domination and Integration Scores in the Paired Comparisons and Teachers' Contacts, Mental Age and Chronological Age of Children in the Combined Morning and Afternoon Sessions of School X ($N = 37$)

			Children's scores		
Teacher's contacts	Teacher	Categories	Domination	Integration	Total interactivity
Total	*A* & *B*	1-24	.39±.09	.15±.11	.22±.11
	A	1-24	.34±.10	.12±.11	.19±.11
	B	1-24	.42±.09	.21±.11	.29±.10
Domination	*A* & *B*	1- 8	.39±.09	.16±.11	.23±.10
	A	1- 8	.33±.10	.16±.11	.22±.11
	B	1- 8	.45±.09	.14±.11	.23±.10
Integration	*A* & *B*	15-23	.30±.10	.04±.11	.10±.11
	A	15-23	.25±.10	—.02±.11	.01±.11
	B	15-23	.35±.10	.24±.10	.30±.10
Mental age			—.44±.09	.17±.11	.22±.11
Chronological age			.09±.11	.10±.11	.12±.11

The child's domination scores show consistently positive and consistently low coefficients of correlation with teachers' contacts. Nine coefficients range from .25 to .45 when the child's domination scores are correlated respectively with the teachers' total contacts, teachers' dominative and teachers' integrative contacts. The coefficients between teachers' integrative contacts and child's domination scores are slightly lower than the other coefficients, revealing differences that in themselves are negligible but that are nevertheless consistent with other analyses of the data on teachers' contacts.

The child's integration scores show coefficients of correlation with teachers' contacts so low as to indicate no relation between the child's integrative behavior in the experimental paired comparisons and teachers' contacts with the child in the schoolroom.

The *child's total interactivity* scores are obtained by combining his domination and integration scores in *the paired comparisons*. The coefficients of correlation represent no relation between the total amount of the child's social interplay in the experimental situation and teachers' dominative, integrative, or total contacts in the schoolroom.

When mental age is correlated with the child's domination scores, the coefficient is —.44±.09. This coefficient indicates for the kindergarten children in this study a consistent and slightly higher relationship as compared with the data in the previous study with *children of preschool age.* There it was found that in one thousand pairings domination scores showed a coefficient of —.27±.02; with 128 pairings when chronological age was held constant the coefficient of correlation between mental age and domination scores was —.35±.05; with 272 pairings among the Iowa City group mental age yielded a coefficient of —.11±.03 when correlated with domination scores (6, p. 389).

With chronological age domination scores show a zero relation. The coefficient of correlation is .09±.11. In the previous study one thousand pairings of preschool children yielded the more definitive zero coefficient of .00±.02. With 128 pairings of preschool children, however, when mental age was held constant the coefficient of correlation between chronological age and domination scores was .24±.06 (6, p. 389).

The integration scores for kindergarten children in the experimental play situation show a coefficient of .17±.11 when correlated with mental age. In the previous study, among 872 pairings a slight positive relation was shown in the coefficient of .30±.03. When chronological age was held constant for 128 pairings a coefficient of .34±.05 was obtained (6, p. 392).

The kindergarten children in the present study show a zero relation between their integrative scores and chronological age. The coefficient of correlation is .10±.11. This is consistent with a zero relation obtained for preschool children on 975 pairings (6, p. 392).

L. THE DYNAMIC NATURE OF INTERACTIVITY AS SHOWN IN DOMINATION AND INTEGRATION SCORES

The dynamic nature of interactivity was shown in Table 26*a and b where the girls who had the lower mean domination scores in* the own-sex pairings showed increased domination scores in the cross-pairings with boys. And the boys who had the higher mean domination scores in the own-sex pairings showed decreased domination scores in the cross-pairings. Although the kindergarten children in this study show a reversal of sex superiority in domination scores obtained in the previous study with preschool children, the dynamic tendency in the cross-pairings confirms the findings of the previous study.

Another way of showing whether or not there is a dynamic relation in domination scores is to correlate the high domination score in each pairing with the domination score of the companion.

TABLE 30

COEFFICIENTS OF CORRELATION BETWEEN HIGH AND LOW SCORES PER PAIRING FOR CHILD'S WITH COMPANION'S DOMINATION SCORES AND INTEGRATION SCORES RESPECTIVELY; COEFFICIENTS OF CORRELATION BETWEEN DOMINATION SCORES AND CHILD'S OWN INTEGRATION SCORES AND BETWEEN DOMINATION SCORES AND COMPANION'S INTEGRATION SCORES

Data on preschool children obtained from Anderson (6).

School	Session	*N*	*r*
	Domination with companion's domination		
X	A.M. & P.M.	63	.85±.02
Y	A.M.	24	.82±.05
Preschool		513	.68±.02
	Integration with companion's integration		
X	A.M. & P.M.	63	.65±.05
Y	A.M.	24	.72±.07
Preschool		514	.82±.01
	Domination with own integration		
X	A.M.	30	.16±.01
	P.M.	96	—.08±.07
	A.M. & P.M.	126	—.01±.06
Y	A.M.	48	.16±.09
Preschool		1,030	—.07±.02
	Domination with companion's integration		
X	A.M.	30	.18±.12
	P.M.	96	—.23±.07
	A.M. & P.M.	126	—.11±.06
Y	A.M.	48	.04±.10
Preschool		1,030	—.10±.02

Table 30 gives the coefficients of correlation of high domination scores with low domination scores for 63 pairings in School *X* and for 24 pairings in School *Y*. The respective coefficients of correlation are .85 and .82. These coefficients show an even higher relationship between domination in one child and domination in the companion than was found in a much larger number of pairings of preschool children. Table 30 includes the previously reported coefficient of .68±.02 obtained from 513 pairings of preschool children (6, p. 395).

Table 30 gives also the coefficients of correlation of high with low integration scores for kindergarten children in Schools *X* and *Y*, together with the coefficient obtained in the previous study with preschool children. Sixty-three pairings in School *X* yielded a coefficient of correlation of .65 and 24 pairings in School *Y* showed a coefficient of .72. These coefficients are lower than, but entirely consistent with, the coefficient of .82±.01 obtained from 514 pairings of preschool children (6, p. 397).

This study thus constitutes additional evidence in support of the findings of the previous study. The hypotheses that domination incites domination and that integrative behavior in one child tends to induce integrative behavior in the companion have received consistent support in both investigations.

Domination scores were also correlated with the child's own integration scores and with the companion's integration scores. Table 30 gives these coefficients for the kindergarten children and reproduces the coefficients obtained previously with preschool children.

Without exception all the coefficients show that a child's domination scores are unrelated either to his own integration scores or to the integration scores of his companion in this experimental situation.

This study therefore offers additional evidence in support of the 15 hypotheses and assumptions which are listed in the introduction to the present study.

M. SUMMARY AND CONCLUSIONS

Aims. The investigation reported herewith constituted a study of dominative and socially integrative behavior of kindergarten children in an experimental play situation. The aims in general were to determine whether methods previously used with children of preschool age would be applicable to kindergarten children;

which hypotheses underlying the previous study of preschool children would find support in the behavior of kindergarten children; which findings of the previous study would not be consistent with the behavior of kindergarten children; *and what relation there might* be between the dominative and integrative contacts of kindergarten children in the experimental play situation and dominative and integrative contacts which the children received from the teachers *in the schoolroom*.

Definitions and hypotheses. Integrative behavior is defined as a phenomenon of growth in which the individual responds voluntarily and without coercion to differences in other persons. In integrative behavior a person yields to another; he finds a common purpose among differences and expends energy with another, i.e., he achieves a change in structure or function, in goals and purposes as a result of encountering persons different from himself. Integrative behavior is spontaneous, dynamic, flexible, changing; it includes the phenomena of spontaneous adaptation and of creative experience or the emergence of originals. Operationally defined, integration is always found at its maximum for the environment as defined. In theory it is growth at the optimum.

Dominative behavior is defined as a technique of responding to others by which a person resists differences, resists change, resists growth. In dominative behavior a person is rigid and inflexible, he has his mind made up; he does not reduce the conflict of differences by finding a common purpose among differences; rather, he maintains or increases conflict or tension between himself and others who differ from himself; he expends energy against or in opposition to others. In dominative behavior a person disregards the desires of others, he uses commands, threats, or force to gain his unyielding objectives; he attacks the status of others; he adds to the insecurity of others.

Domination is something less than spontaneous behavior. It is not yielding or growing; it is self-preserving; it is an expression of fear of an impending change; it is the behavior of an insecure person.

Domination incites resistance or if the balance of power is too great it produces submission. Resistance to domination is itself dominative behavior. Submission, like domination, is a fear response; it is the response of an insecure person afraid of impending change;

it is an effort to preserve a status quo. However justified resistance or submission may be on ethical or social grounds, neither one is related to growth.

There are probably no situations in which one can find domination or integration in a "pure" form. But most situations in which persons find themselves from moment to moment are characterized by one or the other technique of responding to differences.

Subjects. Forty-nine children enrolled in three kindergarten groups were studied.

Procedure. Each child with a few exceptions was paired at random with five others. Two children were brought to the experimental playroom and were allowed to play with a sand box and toys for five minutes, the observer recording their behavior through an observation screen. From the tallies on the observation blanks domination and integration scores were totaled.

Results. The following results were obtained:

1. Coefficients of consistency of two independent observers recording simultaneously were .89±.02 for domination scores and .90±.02 for integration scores.

2. Boys when paired with boys showed higher mean domination scores than did girls when paired with girls; there were 98 chances in 100 that this obtained difference represented a true difference *greater than zero. This was a reversal of the previous findings* with preschool children in which girls were consistently more dominating than boys in the own-sex pairings.

3. In the cross-pairings of the sexes as compared with own-sex pairings the boys' mean domination scores decreased and the girls' mean domination scores increased. In the kindergarten group having the largest numbers of pairings and in the combination of data for two kindergarten groups the increase in mean domination scores for girls approached significance, but the difference in boys' mean domination scores was not significant.

4. In comparing the domination scores of the younger morning kindergarten group with the scores of the older afternoon group there was only one difference that approached significance. When the girls were paired with boys, the older girls had the higher mean domination scores; there were 98 chances in 100 that the difference represented a true difference greater than zero.

5. Girls and boys in own-sex pairings were not significantly different in mean integration scores.

6. In the cross-pairings of the sexes the combined data for two kindergarten groups showed a decrease in mean integration scores for both boys and girls that in neither case was significant. The decreases in mean integration scores for both sexes in the cross-pairings of sexes was consistent with the previous findings with orphanage children in which some of the obtained differences were significant.

7. In comparing mean integration scores of younger children in the morning kindergarten group with older children enrolled in the afternoon group none of the differences was significant. However, when boys were paired with boys, the younger boys had a higher mean integration score than did the older boys, with 90 chances in 100 that the difference represented a true difference greater than zero.

8. In comparing the kindergarten children in the present study with the orphanage children of preschool age in the previous investigation it was noted that the mean domination scores ranged at similar levels. However, the mean integration scores of the kindergarten children for all groups and pairings were in no case less than five times the mean integration scores of the orphanage children.

9. The Domination-Integration ratio of teachers' contacts with these individual kindergarten children showed that domination contacts exceeded integration contacts, the *D-I* ratio of the median child in several groupings being not far from two to one. The *D-I* ratios of children's contacts with each other as shown in mean scores for the groups indicated a great reversal, with integration scores outnumbering domination scores in all cases with one exception by not less than ten to one.

10. The *D-I* ratios obtained from the behavior of orphanage children in the previous study ranging from .54 to 1.19 for the orphanage nursery school group and from .34 to .70 for the orphanage control group are without exception higher than those obtained with the kindergarten children in the present study. The differences are to be explained by the higher integration scores of the kindergarten children.

11. A tendency was found for both domination scores and integration scores to increase from the first to the fifth pairing, the difference for domination scores being significant and for integration scores approaching significance.

12. Nine coefficients of correlation between children's domination scores and teachers' contacts including respectively teachers' dominative, integrative, and total contacts, ranged from .25 to .45.

13. No relation was found between the child's integrative scores in the experimental play situation and the teachers' contacts with the child in the schoolroom.

14. No relation was found between the child's total interactivity scores in the experimental play situation and the teachers' contacts with the child in the schoolroom.

15. A coefficient of —.44±.09 was obtained by correlating domination scores with mental age.

16. The coefficient of correlation between domination scores and chronological age was .09±.11.

17. Integration scores correlated with mental age by a coefficient of .17±.11, and with chronological age by a coefficient of .10±.11.

18. The children's scores for total interactivity correlated with mental age by a coefficient of .17±.11 and with chronological age by a coefficient of .12±.11.

19. When high domination scores were correlated with the companions' domination scores the coefficient for 63 pairings in School *X* was .85±.02, and for 24 pairings in School *Y* .82±.05. These coefficients are consistent with, but higher than, the coefficient of .68±.02 obtained from 513 pairings in the previous study of children of preschool age.

20. When the high integration scores were correlated with the companions' integration scores the coefficient for 63 pairings in School *X* was .65±.05, and for 24 pairings in School *Y* .72±.07. These coefficients are consistent with but slightly lower than the coefficient of .82±.01 obtained from 514 pairings in the previous study of children of preschool age.

21. When domination scores were correlated with the children's own integration scores four coefficients ranged from —.01 to .16, with no probable error less than .06. This zero relationship between domination scores and the child's own integration scores confirms the findings of the previous study in which a coefficient of —.07±.02 was obtained from 1,030 pairings of children of preschool age.

22. When domination scores were correlated with the companions' integration scores four coefficients ranged from —.23 to .18, with no probable errors less than .06. This zero relationship

between a child's domination score and the companion's integrative behavior confirms the findings of the previous study in which a coefficient of —.10±.02 was obtained from 1,030 pairings of children of preschool age.

In a previous investigation of children of preschool age considerable evidence was advanced in support of the following hypotheses:

1. Domination incites domination (resistance).
2. Domination is dynamically related to dominative behavior in a companion.
3. An insecure child makes a companion insecure.
4. Energy expended against a companion will induce energy in the companion directed against oneself.
5. Domination does not induce integrative behavior in a companion.
6. Domination is not only different from, but where a potential avenue of escape is left open it is dynamically unrelated to integrative behavior.
7. An insecure child does not make a companion secure.
8. Energy expended against a companion does not induce energy in the companion directed in a common purpose with oneself.
9. Integrative behavior in a child induces integrative behavior in the companion.
10. Integrative behavior in a child is dynamically related to integrative behavior in a companion.
11. A secure child makes for security in a companion.
12. Energy expended in a common purpose with a companion induces energy in the companion directed with one's own purposes.
13. Integrative behavior does not induce dominative behavior in a companion.
14. Security in a child does not induce insecurity in a companion.
15. Energy spent in a common purpose with a companion does not induce the expenditure of energy in the companion directed in opposition to oneself.

The present study of kindergarten children, while showing some departures from the findings with preschool children, offers only consistent evidence in support of the above dynamic assumptions regarding human behavior.

REFERENCES

1. ALLEN, F. H. The Child as the Therapist Sees Him. *Child Wel. Pamph., State Univ. Iowa,* 1935, No. 50. (Pp. 8.)
2. ANDERSON, H. H. Conflicts in personality development. *Ment. Hygiene,* 1936, **20**, No. 4. (Pp. 605-613.)
3. ———. Dynamic Nature of Personality. *Nat. Elem. Prin., Fifteenth Yearbook,* 1936, **15**, No. 6. (Pp. 244-253.)
4. ———. Children in the Family. New York: D. Appleton-Century, 1937. Pp. xii+253.
5. ———. An experimental study of dominative and integrative behavior in children of preschool age. *J. Soc. Psychol.,* 1937, **8**, 335-345.
6. ———. Domination and integration in the social behavior of young children in an experimental play situation. *Genet. Psychol. Monog.,* 1937, **19**, 341-408.
7. ———. Trends in Mental Hygiene. (In *Marriage*. M. Jung, Ed.) New York: Crofts.
8. ———. Mental hygiene and the wholesome growth of the school child. (In *Mental Hygiene in Modern Education,* Witty, P. A., Skinner, C. E., *et al.*) New York: Farrar & Rinehart, 1939. Pp. 163-183.)
9. FISHER, R. A. Statistical Methods for Research Workers. (5th ed.) London: Oliver & Boyd, 1934.
10. GARRETT, H. E. Statistics in Psychology and Education. New York: Longmans Green, 1926. Pp. xiii+317.
11. ODELL, C. W. Statistical Method in Education. New York: D. Appleton-Century, 1935. Pp. 457.

Published as a separate and in *Genetic Psychology Monographs*, 1939, 21 387-457.

THE CAPACITY OF THE RHESUS AND CEBUS MONKEY AND THE GIBBON TO ACQUIRE DIFFERENTIAL RESPONSE TO COMPLEX VISUAL STIMULI*

William Egleston Galt

Animal Laboratory, Department of Psychology, Columbia University

Acknowledgments 389

I. Introduction 391

II. Method and procedure 401
  A. Subjects 401
  B. Apparatus 404
  C. Adjustment to test situation 412
  D. Techniques employed in testing 413
  E. Experimental controls 426

III. Results 429
  A. Series I 429
  B. Series II 433
  C. Series III 436
  D. Series IV 437
  E. Supplementary analyses 438

IV. Discussion of results 445
  A. Comparison of rhesus and cebus monkey 445
  B. Comparison of monkey and gibbon 448

V. Summary and conclusions 453

References 455

---

*Accepted for publication by C. J. Warden of the Editorial Board, received in the Editorial Office on April 15, 1939, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

ACKNOWLEDGMENTS

The experiment was conducted under the supervision of Professor C. J. Warden of the Department of Psychology of Columbia University and the writer wishes to express his sincere appreciation to him for much helpful advice in regard to the experimentation and for his thoughtful criticisms and suggestions in the preparation of the manuscript. Thanks are also due to Professor Henry W. Nissen of the Laboratories of Comparative Psychobiology, Yale University, for his kindness in allowing the writer to examine an unpublished manuscript of a study covering similar work on the Chimpanzee. The author is especially grateful to the Scientific Director of The Lifwynn Foundation of New York City for generously granting him half-time leave of absence in order to undertake the study here reported.

WILLIAM EGLESTON GALT

LABORATORIES OF COMPARATIVE PSYCHOLOGY
COLUMBIA UNIVERSITY
NEW YORK CITY

I. INTRODUCTION

Although considerable experimental work has been done in the last half century on various types of problems in animals, few of these studies have entailed the comparison of the behavior-capacities of different species. Among the more interesting performances of animals capable of comparison are those relating to the "higher" mental or symbolic processes—processes which have been so greatly developed in man. In dealing with this type of function the primates are exceptionally well suited for study. On the one hand their mental organization is much less complex than that of man and therefore more readily permits of analysis. On the other, they are sufficiently close to man from the point of view of evolutionary kinship, of structural organization and of functional capacity, to make comparisons of scientific value possible.

From the viewpoint of human psychology we are primarily interested in finding indication as low in the evolutionary scale as possible of those symbolic capacities which have been so markedly developed in man. We are interested in determining the limits of such abilities as abstraction and generalization in lower animals, as well as in determining their ability in the performance of such functions as brightness-discrimination and color-constancy. We are further interested in comparing and contrasting such ability with the ability which exists higher in the evolutionary scale and in man. The function of abstraction that is being tested for in this experimental study may be said to be somewhat analogous to this function as it occurs in the developing child. It represents the capacity to isolate specific cues within a complex, and similar analytic procedures.

Nissen (40, 41) has recently made an experimental study on the chimpanzee (four in number) in which he trained the animal in four discriminatory reactions, established successively and cumulatively, so that finally they could be performed in random order. Nissen points out that the experiment tests for three phases or levels of discriminatory behavior: (*a*) response to relation, (*b*) response to absolute properties, and (*c*) response to ambivalent stimuli in which the cues run counter to the reactions of the first two stages. To quote:

> "Absolute properties are no more an indication of the correct response than are constant relations such as larger,

> *smaller, brighter or darker.* Each stimulus is ambivalent, functioning to evoke a positive response in one combination, a negative response in another." He continues: "Response to ambivalent stimuli appears to be a relatively basic category of behavior, extremely important . . . in human language responses, and it should, therefore, fill a gap in analyzing intra-species comparisons and differentiations of behavior."

As will be indicated later, the present study may be regarded as an extension of the general problem to monkeys and the gibbon.

For purposes of general orientation it seems desirable briefly to review such studies of monkeys and apes in which definite comparisons of (*A*) neurological structure and of (*B*) behavioral level have been made. The following types of studies dealing with behavioral level will be discussed: (1) learning, (2) delayed response, (3) abstraction, (4) instrumentation. For more detailed information concerning experimental work on primates, refer to the works of Warden, Jenkins and Warner (56), Yerkes (60), and Zuckerman (63).

A. NEUROLOGICAL STRUCTURE

Students are in pretty close agreement in regard to neurological data on the cortical development of the animal types under discussion. As indicated by Yerkes (60), Zuckerman (63), and Tilney (48), the brain of the chimpanzee is much more complex and resembles more closely that of man than do the brains of the gibbon and the monkey. Tilney finds that the brain of the gibbon has more convolutions, that the super-brain departments for sight, for hearing, and for contact-sense are all more extensive, and that the frontal lobe is more developed than in monkeys but much less so than in the orang-outang or the chimpanzee. He reaches the conclusion that the brain of the gibbon resembles much more closely the structural characteristics of the monkey brain than it does the brain of the higher apes. In his concluding paragraph on the gibbon, Yerkes (60) states:

> We are now entirely convinced by psychobiological facts as well as by the morphological considerations of conventional taxonomy, that these primates . . . constitute a most interesting and distinctive group, consisting of the divergent types gibbon and siamang, between the Old World monkeys and the three existent types of man-like ape.

In comparing the brains of Old World and New World monkeys, Leche (35) gives tables showing that the average brain weight of the cebus monkey is 69.5 grams while its average body weight is 1290 grams. Tilney (48) and Schultz give the brain and body weight of the rhesus monkey as 126 grams and 4310 grams respectively. This gives a brain-body ratio of 1:18.5 for the cebus and 1:34.2 for the rhesus. Thus, while the absolute size and weight of the brain is greater in the catarrhine than in the platyrhine group (126 grams as compared with 69.5 grams), the ratio of brain weight to body weight is much larger for the platyrhine monkey (1:18.5 as compared with 1:34.2). The average brain weight of the gibbon as given by Leche is 130 grams, thus being greater in absolute magnitude than that of either the macacus or cebus monkey. Its brain-body ratio, however, is only 1:73, thus giving it a position considerably below that of the rhesus and greatly below that of the cebus monkey. It is doubtful, however, whether a direct comparison of brain-body ratio is justified between types as divergent as monkey and gibbon.

Zuckerman (63) lists 14 functional characteristics of man, and then indicates which of these features each of the primate groups has in common with him. From this table one gets a good picture of the biological affinity within these groups. In common with man the chimpanzee has all these key features. The only difference is that the chimpanzee has one character which man does not have, namely, the presence of sexual skin. The gibbon corresponds to man in ten of the features, while three have yet to be investigated; the Old World monkeys correspond in eleven features, while the New World monkeys correspond in only eight of the characters. Since two of the features yet to be studied in the gibbon (color vision, related red cells) are almost certain to show correspondence, it can be said with a fair degree of certainty that, ranging in respect of biological affinity with man from most to least, the order is chimpanzee, gibbon, Old World monkey, New World monkey.

B. BEHAVIOR LEVEL

When, however, it is question of the comparative ability in behavioral functions of the primate types under discussion, even in scientific circles there is much diversity of opinion. On the one hand Yerkes (60) says:

> " . . . we risk the prophetic surmise that in respect to the forms of behavioral adaptation to captivity in point, the difference between chimpanzee and man is less than between chimpanzee and Macacus monkey, or between the latter and the lemur."

On the other hand, Zuckerman (63) says:

> "What is important is that the results of such experiments on primate behavior as have been carried out up to the present offer little or no support to the a priori belief that apes are more intelligent than monkeys."

The views of other investigators range between these two extremes.

1. *Learning*

Haggerty (21) tested the comparative ability of ten cebus and two rhesus monkeys on a series of problem boxes and came to the conclusion that neither species had the advantage over the other. Koch (33) directly compared the limits of learning ability in the cebus monkey with the limits determined by Dr. Fjeld (10) for the rhesus monkey. Koch used six animals and Fjeld seventeen, and they both made use of the same learning apparatus (the Jenkins problem box). Conditions for the two studies were approximated as closely as possible, so that valid comparisons could be drawn. Koch found that there was no significant difference between the limits of learning in the two species, but that the cebus monkeys learned the different steps in approximately one-half the number of trials required by the rhesus. Boutan (5) found that the one gibbon tested by him could open a variety of problem boxes and that its retention was perfect for these performances over a three-month period. He reports that the gibbon fatigued easily when its attention was strongly held by visual or auditory stimuli, or when novel problems were presented to it, and that on these occasions it became restless and irritable.

2. *Delayed Response*

Harlow, Uehling and Maslow (22) found that the New World monkey was no better than the lemur in delayed reaction tests, while the Old World monkey was definitely superior to them both. In a further study Harlow (23) found that certain of his Old World monkeys could solve two delayed reaction problems that were being carried on concurrently, while none of the New World monkeys tested could do so. The one gibbon tested was on the same level as the rhesus monkeys in the simple delayed reaction tests and was inferior to the best of the Macaques in complicated delayed reactions. The animal was difficult to motivate. On the basis of his studies on delayed responses Harlow and his collaborators (22, 23) ranked the types of animals under discussion in the following order of ability: anthropoid apes, Old World monkeys and gibbon, New World monkeys.

In a study comparing the ability of the monkey and chimpanzee in multiple delayed reactions, Tinklepaugh (49) found that after an eight hour delay his two monkeys were making scores of 56 per cent and 66 per cent correct (50 per cent representing chance performance) whereas the two chimpanzees tested were correct in 85 per cent and 90 per cent of their choices. It is interesting in this connection that Jacobsen (27, 28) has found that extirpation of the frontal association areas (the areas preeminently mediating symbolic processes in man) profoundly disturbs the delayed response of monkeys and chimpanzees, whereas it does not disturb such functions as visual discrimination, manipulation of problem boxes, etc. This same investigator found that bilateral removal of other portions of the brain did not interfere with the delayed response, and these results are in harmony with those of Breslaw, Barrera and Warden (6), who found that the ability of their monkeys in the delayed response was not decreased but rather augmented by the bilateral extirpation of the post-central convolution. The studies here cited are in general agreement with the extirpation work that has been done by other investigators on the primate. There would, therefore, seem to be ample ground for considering the delayed response as dependent upon the symbolic function and as a good index of ability in the "higher" mental processes.

3. *Abstraction*

Gellermann (16) found that both chimpanzees and two-year-old children could discriminate form *per se*, although children were superior, and that both gave evidence of symbolic behavior associated with the discrimination of form. Neet (38), in testing four macacus rhesus monkeys on visual discrimination patterns, found that one of the animals was able to respond to triangularity or form *per se* while the other three were unable to do so. He disagrees, however, with Washburn's contention that such a response is due to the presence of "an abstract idea of triangularity." Yoshioka (61) was able to train four young chimpanzees always to respond to the "odd" one of three food containers, and to choose three food cans in the temporal order 1, 2, 3, in three consecutive trials, repeated five times in a single setting. Yoshioka therefore holds that the chimpanzee is capable of *"inhaltlich"* if not *"formale"* abstraction. Robinson (45) in an exploratory experiment taught one male *Macacus cynomologus* always to respond to the "odd" one of three stimuli. Gellermann (14, 15) found that his rhesus monkeys could learn the double alternation of four responses in the temporal maze and could extend the double alternation of responses beyond the length of the training series. In accordance with the position of Hunter, he interprets his results as indicating the presence of symbolic processes in his subjects. Gallis (12) was able, after a long period of training, to establish in two young monkeys (*Macacus sinicus*), the "idea" of the number three, while Bierens de Haan (3) was unable to get positive results on a similar experiment with a two-three year old Java ape, and concluded that the ape has no conception of numbers.

It might be mentioned at this point that when working with one or two animals, as is the case in the majority of studies here reported, positive results have much greater weight than negative ones. Negative results may be due to the fact that the one or two animals being tested are "dull" or below par in performance, and do not necessarily indicate that a typical member of the species would not be able to perform the task. Valid positive results, on the other hand, indicate that the species has reached a level of development that permits the performance of the task in question.

Tellier (50) found that a young female monkey (*Macacus rhesus*) was able to learn that the same box would never contain food twice in succession, and that a young *Macacus sinicus* successfully acquired the habit of always responding to "the second from the left" when the number of containers in the row to choose from was changed from trial to trial. As a result of experimentation on a bonnet-macaque monkey, Verlaine (51, 52) concluded that the performance of his monkey proved "the existence of powers of analysis and synthesis in the macaque which are singularly like those which condition the so-called higher processes of abstraction."

4. *Instrumentation*

Harlow and Settlage (24) found that cebus monkeys were slightly inferior to rhesus monkeys in patterned string tests but that individual differences within a species were also great. Experiments by Nellmann and Trendelenburg (39) indicate that in drawing food toward themselves with a rake the monkey and the gibbon show about the same level of adaptive ability. Drescher and Trendelenburg (9) and Guillaume and Myerson (18) found that the gibbons tested by them did not show much ability in instrumentation. They say: "once more we establish here that these anthropomorphes scarcely arrive, from an intellectual point of view, at the level of the lower monkeys." Two gibbons failed completely on the diagonal cord tests of Guillaume and Myerson (20), although monkeys usually succeeded in solving this problem. Chimpanzees and a gorilla, on the other hand, showed a marked superiority over the monkey on complicated string tests.

Verlaine and Gallis (55) tested a young *Macacus sinicus* on problems of a similar nature to those given by Köhler (34) to chimpanzees and found that it behaved as intelligently as Köhler's anthropoids. It not only used objects as tools but fashioned a tool for itself, was successful in the detour experiment, the stool experiment, and that of the obstacle to be avoided. It should be mentioned, however, that other experimenters have not found monkeys so proficient in these experiments. Guillaume and Myerson (18) found that their gibbon, like their monkeys and higher anthropoid apes, was not deterred from pulling a string, to which food was attached, when the food did not begin to move immediately, or *even* when it moved temporarily in the opposite direction (detour problem). Most investigators have not found this type of response in the monkey and have looked upon it as characteristic of anthropoid apes and of man (Cf. Yerkes, 60).

Bierens de Haan (3) is the only investigator who has obtained positive evidence of box-stacking in monkeys. His cebus monkey not only succeeded in stacking two boxes but learned to collect three boxes when they were put in different corners of the cage, and to stack them on top of one another. This performance compares favorably with performances on this problem by chimpanzees. Harlow and Settlage (24) found that their monkeys did as well on the patterned string test as Köhler's chimpanzees had done, but remark "it is possible that extensive researches would reveal the superiority of the anthropoids."

Yerkes stresses the facile use of tools by the chimpanzee as indicating the wide gap in mental ability between it and the monkey, but experimental studies in the last years by Bierens de Haan (3), Verlaine and Gallis (55), and Klüver (31) show that at least the cebus monkey has considerable ability in the manipulation of instruments. A very recent study by Warden, Fjeld and Koch (as yet unpublished) compared the ability of the rhesus and cebus monkeys in instrumentation and found that the cebus was not only superior to the rhesus but that its performance compared favorably with that of the chimpanzee. In the single platform set-up the monkeys were brought to the same level of performance as the chimpanzees in the unpublished work of Jackson. It is only fair to point out that the limits were not tested for in either study, and to state that the chimpanzees required much less training than the monkeys.

It is doubtful, then, if there is such discontinuity in behavior-performance between the monkey and the anthropoid as is posited by Yerkes. Recent studies have tended to show that the monkey is capable of more complicated and more ape-like or man-like responses than it was formerly credited with. While earlier experimental findings favored the capacity of the rhesus over that of the cebus monkey, later investigations have found the cebus to be superior in a number of performances. What scant experimental work there is on the gibbon indicates that these animals are certainly not superior, and perhaps may be somewhat inferior, to monkeys in solving the type of problem ordinarily required of primates in the laboratory.

5. *The Problem*

As indicated above, the present problem was similar in general to the work of Nissen on chimpanzees. In the usual discrimination experiment the stimuli employed vary only in a single dimension, such as size, brightness, form, and so forth. The work of Nissen and that of the present experiment goes beyond this in that the sets of stimuli used involve both brightness and size. In addition to this greater complexity, the problem was further increased in difficulty by the presence of two factors: (*a*) Antagonistic cues within each series of stimuli, and (*b*) ambivalent cues between the several series.

These relationships are indicated diagramatically in Figure 3. The aim was to determine, if possible, the limits of complexity which could be mastered by the cebus monkey (*Cebus capucina*), the rhesus monkey (*Macacus mulatta*), and the gibbon (*Hylobates lar*). This would enable us to compare the capacity of the several types with the ability of the chimpanzee as shown in the results secured by Nissen.

If the present study can demonstrate that the monkey is capable of the very complex perceptual and "higher" mental processes entailed in the cumulative establishment of the four discriminatory responses called for by our experiment, it will afford added indication of the close mental continuity existing between the monkey and the higher anthropoids. If, on the other hand, it is not possible to set up this complex series of habits in the monkey, it will be rather strong indication of a fundamental difference between it and the chimpanzee in complexity of mental organization. The study should also indicate the comparative ability of the rhesus monkey, the cebus monkey, and the gibbon, to execute the complex type of perceptual response demanded of them.

II. METHOD AND PROCEDURE

The experimental work to be reported was carried on in the laboratory of Comparative Psychology, Columbia University, from December 1, 1937, to June 1, 1938, and consisted of training monkeys in four antagonistic visual discrimination-habits. Experimentation was conducted in a specially built dark room. This room was located in a quiet portion of the building and was fitted with double doors leading into the hallway. Although the room was not completely sound-proof, these doors lessened the noises coming from the outside. The experimental room was some little distance from the primate vivarium so that the animals while working in the apparatus were never disturbed or distracted by noises from the other monkeys in their living-cages. A rubber-tired truck was used to transport the monkeys from their living-cages to the experimental apparatus. The gibbon was carried in the arms of the experimenter as it was gentle and as this constituted for it much the more natural mode of transportation. The animal subjects were kept under uniform conditions throughout the experiment.

This section may be conveniently treated under the following four headings: (*a*) Subjects, (*b*) apparatus, (*c*) adjustment to test situation, (*d*) techniques employed in testing.

A. SUBJECTS

The subjects in this experiment were three young rhesus monkeys (*Macacus mulatta*), two young cebus monkeys (*Cebus capucina*) and one young gibbon (*Hylobates lar*). None of these animals had previously been trained on discrimination problems. The monkeys were all purchased from an animal dealer and had been at the Columbia Laboratory for two years. They had been used on a series of laboratory problems—pulling in strings, instrumentation, and imitation tests—over a period of about a year and a half. For six months prior to the present experimentation they had not been used on any problem. The gibbon was loaned by Dr. C. R. Carpenter of Bard College and had been in the primate vivarium at Columbia for six months before the beginning of the present study. It had not been used in any previous experimentation.

Protocols

Rhesus No. 1. Female. Weight 156 ounces. Estimated age five years. This animal was moderately tame throughout the experiment. She was somewhat timid and very cautious in her reactions during the first part of the experiment. At times she did not appear well motivated. As work progressed, she seemed to give more attention to the stimuli and to be more interested in the reward. The animal was not restless during the experimental sessions. Between discriminations she assumed a characteristic position, sitting quietly on the adjustable platform to the right of the shutter-opening with her face sideways to the opening and with her hands crossed.

Rhesus No. 2. Male. Weight 140 ounces. Estimated age four years. This animal was tame but was the most active of the subjects. He seemed always very restless and between discriminations was constantly running about in the reaction-cage and biting at the sides of the cage and at the adjustable platform. For the most part he was attentive while making discriminations and showed more comparison-behavior than did the other two rhesus monkeys. In the late stages of experimentation he seemed to respond in a more haphazard manner.

Rhesus No. 3. Female. Weight 160 ounces. Estimated age five years. This animal was quite tame. She was intermediate in activity between Rhesus No. 1 and No. 2. She was quite restless during the first part of the experimentation but later became very quiet in the experimental situation. One characteristic of the animal was that she spent the time between discriminations in picking at herself or scratching. For the most part she seemed to give attention to the stimuli and appeared to be well motivated.

Cebus No. 1. Male. Weight 52 ounces. Estimated age four to four and one-half years. This animal was very tame and liked to be handled. He was most attentive during discriminations, seemed to compare the stimuli carefully and, after learning the first discrimination, progressed rapidly. In the early stages of the experiment, his right hand was badly ripped by one of the rhesus monkeys and he had to be sent to the Bronx Zoo Animal Hospital for two weeks. The wound healed nicely, leaving no manipulatory impairment and no apparent emotional disturbance. As the accident occurred before the animal had learned his first discrimination and as his score on this discrimination was as good on his return to the problem as it was prior to the accident, it is believed that the interruption caused no serious disturbance in the learning-process. This animal appeared well motivated throughout the experiment. Towards the end of the work, raisins and small cubes of apple were substituted for peanuts as the food-reward as the animal seemed to tire of the latter.

Cebus No. 2. Male. Weight 72 ounces. Estimated age four and one-half to five years. This animal was for the most part very tame. He was somewhat more restless and active than Cebus No. 1, became disturbed more easily over errors, bit at anything within reach and was destructive at such times. The animal liked to be handled by the experimenter and during part of the period of experimentation had to be conveyed manually to the transfer-cage and to the reaction-cage. He was especially fond of going through the pockets of the experimenter. This animal was very attentive during discriminations, showed much comparison-behavior and was well motivated throughout the experiment. Raisins and small cubes of apple were alternated with peanuts as the food-reward during the closing days of work. A mild experimental neurosis was set up in this animal.

Gibbon ("Lo"). Female. Weight 116 ounces. Estimated age three and one-half to four years. This animal was very tame and affectionate. She liked to play with the experimenter and to jump on him and put her arms about his neck, and was in this way carried to and fro between the living quarters and the experimental room. She enjoyed swinging about in the exercise-cage and became cross and dissatisfied if not allowed to do so. The animal seemed well motivated throughout the experiment. She was a good and tireless worker. She was intermediate between the rhesus and cebus monkeys in the degree in which she showed comparison-behavior. At times she exhibited typical "temper tantrums".

The general program of care and diet of the experimental animals is one that has been standardized in the Columbia Laboratories and described in detail by Fjeld (10), Koch (33), and Crawford (8), and so need not be discussed here. The daily routine consisted of two periods of experimentation for each animal. Each period lasted for about 20 minutes and each animal had approximately two hours rest between periods. After the first work period, each animal was given a half pint of milk and eggs (two eggs to a quart of whole milk). As soon as all the animals had finished their second work period, at noon or shortly after, they were fed their chief meal. They then received no more food until the milk and egg mixture the following morning.

As food-reward during the training period, the rhesus monkeys were given one small raisin on making a correct choice, the cebus monkeys a quarter of a peanut, and the gibbon a small cube of apple. None of these foods were included in the regular diet of the animals and they produced a good degree of motivation throughout the experiment. In the closing stages, the two cebus monkeys seemed to tire of peanuts and these were varied with raisin and apple. When making 100 per cent or near 100 per cent correct responses, the cebus monkeys appeared to get more peanuts than they wanted and the reward lost its motivating power to some extent. But as I shall emphasize more in detail later, these animals seemed to be impelled by the work itself and would keep on making the discriminations even though not eating the peanuts after securing them. It hardly seemed practical to reduce the reward to less than a quarter of a peanut. The rhesus monkeys were never satiated with raisins, and the gibbon was always eager for cubes of apple. The cubes were of such a size that 50 could be cut from one-half of a medium-sized fruit, so that the animal never secured more than a half of an apple each day even though all its choices were correct. For the most part, of course, it secured much less than that.

At the beginning of experimentation all of the animals except the gibbon were fitted with leather belts to each of which was attached a light chain. Early in the training, the chain was removed from Rhesus No. 1, as she tended to hold it in her hand while in the reaction-cage, and this interfered with her response. Some weeks later the belt and chain were removed from Rhesus No. 3, as the former was rubbing a sore spot on her side.

B. APPARATUS

The apparatus used in the two methods (reaching and pulling-in) will be described in order. The reaching technique involved the use of a reaction-cage, a stimulus-incentive box, and a table. Figure 1 shows a diagram of these three parts as they were arranged in the experimental sessions. The dimensions of the reaction-cage were 77.5 cms. x 69.2 cms. x 78.1 cms. A 1.27 cm. wire mesh, 16 gauge, covered the top of the cage, the side toward the experimenter, and

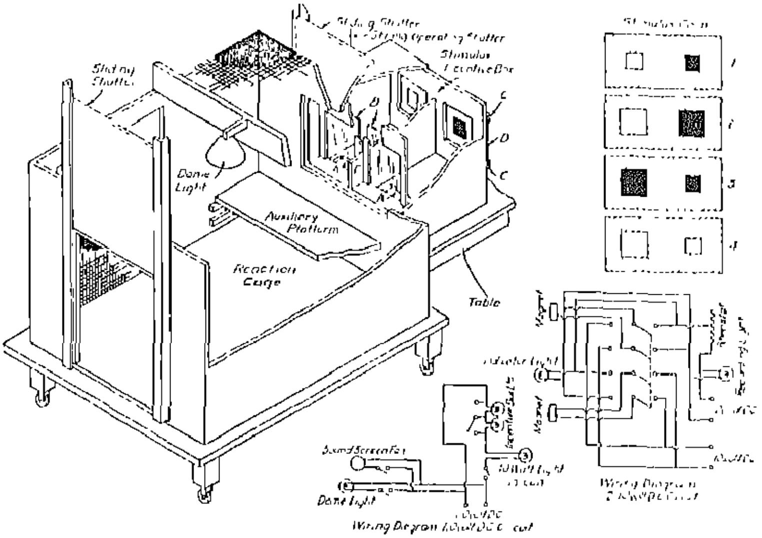


FIGURE 1[4]

the lower part of the door. Window screening, painted white, covered the mesh. The remainder of the cage was covered with a .24 cm. planking. The screening acted as a one-way light screen when lights were out in the experimental room and the dome-light in the reaction-cage turned on. The cage was supported upon a platform 34.3 cms. high, having a top surface 124.3 cms. long and 76.2 cms. wide.

The opening through which the animal saw the stimulus cards and through which it reached when opening the glass doors of the *stimulus-incentive box* was 27 cms. wide by 22 cms. high and was in the forward wall of the cage. This opening was closed by means of a sliding door operated by a cord running over a pulley above. In the rear cage-wall there was located another door swinging outward, into which was fitted a smaller vertically sliding door. This latter door was on a level with a similar sliding door of the transfer cage, so that when both doors were raised the animal could jump directly from transfer-cage to reaction-cage and vice-versa without any possibility of escaping into the room and without having to be handled.

---

[4] *A*, food wells; *B*, Vita glass doors; *C*, grooves for stimulus cards; *D*, edge of stimulus card; *E*, partition between pair of stimuli.

The reaction-cage was lighted by a dome-light which consisted of a 25 watt frosted globe in a 15.2 cms. desk-lamp reflector. There was an auxiliary platform in the bottom of the cage which was adjustable to different heights in order to give the differently sized animals the correct position in relation to the front opening. The platform and inside of the reaction-cage were painted a flat white, the outside a battle-ship gray.

The stimulus-incentive box was a device designed and used by Crawford as an incentive box and described in detail by him (8). Essentially it was an oblong box, 43.2 cms. wide, 12 cms. deep, and 27.3 cms. high with a .63 cm. thick partition in the center dividing it into two compartments each 21.3 cms. wide and 12 cms. deep. The box was fitted with a sliding tray into which two food wells were so cut that when the tray was in place, one well was in each compartment. No light could pass from one compartment to the other when the tray was in position. In the front and back of the stimulus-incentive box were two openings 16 cms. wide and 12.5 cms. high. The front pair of openings was fitted with *Vita Glass* doors hinged at the top and opening inward. Vita glass was used as the coefficient of refraction is so much smaller than it is for ordinary glass. These two glass doors were connected in such a way with magnets, actuated by currents of two and ten volts, that one door was just held closed whereas the other door was locked. The door to be locked could be alternated from side to side by a single movement of a double-pole, double-throw switch which was fastened to the top of the box. It was found necessary to place a rheostat in the two-volt circuit and thus further reduce the current, as the unlocked door was held closed with sufficient force to discourage some of the animals (especially the gibbon) from opening it. A small two-volt bulb whose glow was scarcely visible showed that current was passing through the two-volt circuit, and it could be estimated from this light whether the current remained practically constant.

The rear pair of openings in the stimulus-incentive box framed the stimulus patches. Two metal grooves were placed on the back of the box, one above and one below the openings, and the stimulus-cards, each of which bore two stimulus-patches, were so constructed that when slid into these grooves, the stimulus-patches were centered in the openings. In order to change a stimulus-patch from right to left or vice-versa it was merely necessary to remove the card, reverse its position, and re-insert it into the grooves. The stimulus-patches were in very close proximity to the food-reward, being only 11 cms. distant and directly behind it. The response was thus in close proximity to the stimuli and the set-up was therefore of the approved type, since other experimenters have amply demonstrated that proximity of stimulus, response, and reward is important in the establishment of an animal's discriminatory reactions.

One small light of six to eight volts was placed in the center of the top of each side of the stimulus-incentive box in order to light the stimuli. These lights were so placed that they could not be seen by the animal. They were connected in series with a 40 watt lamp which was concealed in a tin container on the floor. From time to time the small bulbs burned out and had to be replaced. Care was taken to put new bulbs in both sides at the same time in order to insure that the same amount of light was given out by each. At the beginning of the experimentation a strip of black paper was pasted over the bottom of each glass door of the stimulus-incentive box in order to insure the animals' not seeing the food in the food-wells. Subsequent experimentation, however, showed that this precaution was unnecessary, and these strips were later removed as they could possibly have interfered with the animals' view of the larger stimulus-patches. In any event both wells of the food tray were baited in each trial, so that no differential cue could have been obtained even though it were possible for the animal to see the food. After one of the cebus monkeys had broken a glass door by pushing it forcibly and thus hitting it against the back of the stimulus-incentive box, sponge rubber pads were placed at these points to help absorb the shock.

Extending vertically outward from the narrow strip that separated the two glass doors on the face of the incentive box was a wooden partition. This partition, which projected 7.9 cms. out from the box, was 15.2 cms. high and 1.96 cms. wide. Also projecting 7.9 cms. outward from the sides, top and bottom of the front of the box was a piece of 1.3 cm. 16 gauge wire-mesh. As, notwithstanding this precaution, the animals tended to peer out between the reaction-cage and the stimulus-incentive box, the spaces below the doors where the magnets were located were boarded up. This prevented the animals from seeing out below and thus helped to center their attention on the lighted stimulus-cards.

The stimulus-incentive box was placed on a table so that it was brought to the proper level directly in front of the sliding door in the forward wall of the reaction-cage. The animals could thus comfortably reach into the incentive box. The switches used for controlling the various electrical appliances were located on the table, as was a two volt light used for purposes of recording. This light was screened on the side toward the animal by black cardboard.

As shown in the schematic wiring diagram in Figure 1, two electrical power circuits were used. The 1/110-v. *DC* circuit was used for the small bulbs in the stimulus-incentive box which lighted the stimulus-patches. These were connected in series with a 40 watt bulb which was concealed in a container under the table. A switch for turning on the two small lights was placed on top of the stimulus-incentive box, and a double throw switch, which made it possible to light the bulb in either side of the box independently, was in a convenient location on the control table. The bulb in the dome-light of the reaction-cage and the electric fan were also operated by this current, as were the two reflector lamps used for lighting the stimuli in the pulling-in method. The 2-10-v. direct current obtained from storage batteries was used for actuating the locking magnets on the stimulus-incentive box, the indicator-light, and the recording-light. Two and ten volt circuits were connected to the arms of a double pole, double throw switch whose opposite binding posts were connected to the two magnets on the stimulus-incentive box. In one position of the switch, M_1 would have a two-volt current flowing through it and M_2 a ten-volt current, and in the opposite position of the switch the currents in the two magnets would be reversed. This arrangement served to close one door and to lock the other with a single movement.

The experimental room was dark except for such light as came from the dome-light of the reaction-cage, the two small lights in the incentive box, the small recording light and, when the pulling-in technique was employed, the two reflector lights that lighted the stimulus-cards. The dome-light in the reaction-cage was regularly extinguished during discriminations after the early part of the experiment.

a. Stimuli. The stimuli were cards on which the appropriate designs were pasted. They were in general similar to those employed by Nissen. The cards used in the reaching method were 46.5 cms. long and 16.5 cms. high, were cut from grey Bainbridge Board and, as mentioned, previously, slid into metal grooves attached to the back of the stimulus-incentive box. The two openings in the back of the box over which the cards fitted were 16 cms. long and 12.5 cms. high, so that this represents the amount of stimulus-card exposed to the animal through each glass door. The stimulus-cards had black and white stimulus-patches pasted on them in such a position that one patch was centered in each opening when the card was slid into position. During the training-series the stimulus-patches were of two sizes; the larger ones 9 cms. sq. and the smaller ones 5 cms. sq. These are of the same relative sizes as some of those used by Nissen in his study and are one-half the actual sizes. It was feasible to have our cards smaller in actual size as they were presented at a distance of only about 25 cms. from the animals' eyes, whereas Nissen's stimuli were some 225 cms. distant from his animals.

One difference between our stimuli and those of Nissen might be mentioned here. In Nissen's work, the background for the stimuli was a neutral screen so large that the edges were not visible to the animal. In our own study, the stimuli were pasted on grey cardboard, the visible edges being 16 x 12½ cm. It seems unlikely, however, that this difference was of much importance. In the first place, the stimulus-patches were made of cardboard 2 mm. in thickness which made them stand out rather prominently from the background. In the second place, the size of the background was always the same and hence could hardly furnish effective cues. Thirdly, the animals appeared to be responding to the stimulus-patches, and this observation is corroborated by certain critical trials discussed under *Results.*

There were four stimulus-cards to a set and they carried the following stimulus-combinations: (*a*) A 5 cm. sq. black patch and a 5 cm. sq. white patch. This combination will subsequently be called small *b* vs. small *w*, or simply *b* vs. *w*. (*b*) A 9 cm. sq. white patch and a 9 cm. sq. black patch. This combination will subsequently be called large *W* vs. large *B*, or simply *W* vs. *B*. (*c*) A 9 cm. sq. black patch and a 5 cm. sq. black patch. This combination will subsequently be referred to as large *B* vs. small *b*, or simple *B* vs. *b*. (*d*) A 5 cm. sq. white patch and a 9 cm. sq. white patch. This combination will subsequently be referred to as small *w* vs. large *W*, or simply *w* vs. *W*. In the critical series, different sized stimulus-patches were used, and specific mention of this fact and of the sizes employed will be made at these points.

Several sets of stimulus-cards were made up and the sets were used alternately in order to guard against discriminations being set up on the basis of secondary cues. After an animal was making 100 per cent correct responses on a stimulus-combination of one set, it was tested with the corresponding card from another set to see if the responses were still consistent. As the cards of a set became soiled they were discarded and replaced by new ones.

b. Modification of the apparatus for pulling-in technique. For the pulling-in method the same reaction-cage was used, the only alteration being that a wooden grill was placed over the shutter-opening at the front of the cage to prevent the animal from escaping from the cage while the shutter was raised. A platform 180 cms. long, 70 cms. wide, and 7.5 cms. high was built of boards and placed on top of the stimulus-table. At a point on the platform 75 cms. from the shutter-opening of the reaction-cage, wooden forms were attached, and into these forms were slipped wooden blocks. Upon these wooden blocks were glued the stimulus-cards.

The stimulus-cards used in the pulling-in method were identical with those used in the reaching-technique except for modifications necessitated by the altered method of presentation. The cards were fastened to wooden blocks and were 12.5 cms. high and 16 cms. wide, i.e., the same area of card as that visible to the animal in the reaching-method. These cards had stimulus-patches centered on them, and a stimulus-combination consisted of two of the cards with their corresponding patches.

An opening was tunnelled through the bottom of each stimulus block from front to back, and fitting into this opening was a small tray 9 cms. long, 7.3 cms. wide, and 2.5 cms. high. The stimulus-cards were attached to the blocks at a point 3 cms. from the bottom. The tray slid into the stimulus-block just beneath the card, allowing less than 0.5 cm. space between the top of the tray and the bottom of the stimulus card. To the center of the front of the two trays were attached black cords having a diameter of .2 cm., which ran parallel along the board to the end of the platform at the opening of the reaction-cage. At a point 1.5 cms. from the end of the platform (nearest reaction-cage) a wire staple was placed over each cord. These served the dual purpose of keeping the two strings parallel and of preventing the animal from drawing the tray into the reaction-cage. Later in the experiment it was found necessary to place other staples 12 cms. from the end of the platform, as some of the animals (notably the gibbon) would take hold of the string beyond the first staple and thus were able to pull the tray into the cage where they would hold on to it, bite it, and try to pull it apart.

The blocks to which the stimulus-cards were attached were 9 cms. high, 10 cms. wide, and 6 cms. deep. They all fit interchangeably into the forms on the platform so that any stimulus-card could be easily presented on either side merely by inserting the proper block. For the training part of the experiment there were four blocks with the following stimulus-cards attached to them: *W*, *w*, *B*, *b*. This gave all the combinations necessary. When two stimulus-cards were in place there was a distance of 5 cms. between the proximal edges of the cards. The distance between the strings was 17.8 cms. The ends of the strings were knotted in order to facilitate the animals' manipulation of them. When the strings were stretched tight the front of each tray was flush with its corresponding stimulus-card. The weight of the tray which the animal pulled in and into which the food-reward was placed was 42.6 grams. The platform, the blocks, and the food trays were all painted a dull black. A gray screen was placed behind the platform in order to give the animals a neutral background. The stimulus-cards were lighted by two 75 watt frosted bulbs which were mounted in reflectors at a distance of one meter directly in front of and above the cards. The reflectors were adjustable and were trained on the stimulus-cards. Both cards were lighted with equal brightness and no shadows were cast on either.

This arrangement with blocks and trays seemed to us to have several advantages for our particular purposes over the pulling-in set-up usually used where the blocks supporting the stimuli are themselves either pulled in completely or part way by the animal. Our stimulus cards did not move and there could therefore have been no confusion due to apparent change in size of the stimuli as they came nearer to the eyes of the animal, nor were the stimuli in danger of soiling or mutilation resulting from handling by it. Further, the trays that were pulled in were very light and required the expenditure of much less effort on the part of the animal than is the case where the stimulus-bearing blocks are themselves pulled in. Then there was the additional circumstance that the transition from the reaching-method seemed to be more natural under these conditions.

C. ADJUSTMENT TO TEST SITUATION

In order to accustom the animals to the experimenter, the latter let them out singly in the exercise cage and fed them raisins and peanuts. They were later led about on chains and were pulled gently by their chains into the transfer truck. After the animals were accustomed to entering the truck, they were next trained to jump into it as soon as the door to their living cage was opened. They were wheeled about the room in order to get them accustomed to the motion of the truck.

After the animals had been thoroughly accustomed to the experimenter, they were wheeled to the experimental room and placed in the reaction-cage. They were trained to jump immediately from the transfer truck into the reaction-cage as soon as both doors were raised, and to take a standing position on the adjustable platform in front of the sliding door. As all the animals except the gibbon had been used in a similar work-cage previously, they readily entered it. The gibbon was placed in the cage by hand. Before bringing an animal into the experimental room, it was darkened, the dome-light in the reaction-cage turned on, the electric fan started, and the two small bulbs in the incentive box lighted.

The sliding door to the reaction-cage was next raised and the animals were given the food-reward (raisin, peanut, or apple) manually. The food was later placed on the ledge of the incentive box and the animals were accustomed to reach for it there. It was next placed within the incentive box, in the well of the sliding food-tray, while the animals were watching. The glass doors to the incentive box were at first held open by the experimenter, as none of the animals would push open the door. They were gradually accustomed to the motion of the glass door as it closed upon the animal's hand in withdrawing it from the box. This motion and pressure upon their hands was at first very disturbing to the animals, but they quickly became accustomed to it. The gibbon, due to its long and rather awkward hands from a manipulatory viewpoint, had considerable trouble at first in picking up the cube of apple from the well of the food-tray, but after a considerable number of practice periods became adept and quick in this manipulation. This animal was also more timid than any of the others in reaching through the glass doors but completely overcame this inhibition during the period of adjustment. During this adjustment period both doors were unlocked, and a uniform gray card covered both stimulus openings.

D. Techniques Employed in Testing

This section will be discussed under the two headings: (*a*) *Reaching technique* (*b*) *Pulling-in technique.*

1. *Reaching Technique*

After all the animals had become adjusted to the experimenter and to the apparatus, and all necessary adjustments and modifications had been made to the latter, work on the first discrimination, "*b* vs. *w*", was started on all animals. The usual routine was to give each animal 40 trials a day in two shifts of 20 trials each. For the most part experimentation was carried on seven days a week and was always in the morning hours. There was no punishment used for incorrect responses beyond withholding the food-reward. The animals were trained to withdraw immediately into the reaction-cage after making an incorrect choice and finding the door locked. They were never allowed to reach into the correct door after first responding to the incorrect one. It was occasionally necessary during the early days of training to let down the sliding door gently on the head of an animal in order to insure its withdrawing immediately into the cage after an incorrect choice.

At the beginning of each experimental period the apparatus was set in order, the locking magnets tested, the dome-light in the reaction-cage and the small lights in the incentive box turned on, the fan started, and the room darkened. The animal was then wheeled from its living quarters and immediately transferred into the reaction-cage. It was allowed to remain in the reaction-cage for about one minute before the first discrimination was presented.

The following procedure was carried out for each discriminatory response of the subject: (*a*) Both wells of the food-tray in the stimulus-incentive box were baited with the proper food-reward. (*b*) The correct stimulus-card was fitted in the grooves on the back of the stimulus-incentive box. (*c*) The magnet switch was thrown in the proper direction thereby locking one and closing the other glass door of the incentive box. (*d*) The dome-light of the reaction cage was extinguished and the sliding door slowly opened. The turning-off of the dome-light served as a ready signal for the subjects. (*e*) As soon as the shutter was opened the stop-watch was started. (*f*) The general behavior of the animal, the number of times it moved its head in looking from one stimulus patch to the other, the length of time elapsing before its response, and its choice of door was observed. (*g*) As soon as the animal had withdrawn into the reaction-cage the sliding door was let down. (*h*) The dome-light was turned on. (*i*) The magnet switch was opened. (*j*) Observations were recorded by the experimenter.

About 45 seconds elapsed between the beginning of one discrimination-trial and the next. Thus a session of 20 trials consumed about 15 minutes. As the animal had two sessions a morning it was at work for approximately a half-hour each day.

Data were taken on blank forms on which spaces were provided for the following data: (*a*) Stimulus-combination, (*b*) position, (*c*) discrimination time, (*d*) score, (*e*) comparisons and (*f*) remarks. The position column was filled in at the time of mimeographing and was the same on all sheets. It consisted of a series of right (*R*) and left (*L*) notations arranged in random order in accordance with the rules laid down by Gellermann (13) for obtaining chance successes. Each series of ten trials contained five rights and five lefts; at least two rights and two lefts appeared in both the first and last half of each series; each series contained only five reversals from right to left or from left to right; and no more than three rights or three lefts ever appeared in succession. The series thus offered a chance score of 50 per cent correct from either simple or double alternation of response. The *R*'s and *L*'s were to designate the first named of a stimulus pair. Recording was never started at the same place in a record sheet on two successive experimental periods, so that there was continual variation of the position of sequence.

In half of the record sheets the "Stimulus" column was left blank and in the remaining half it was filled in with "*b* vs. *w*" and "*W* vs. *B*" designations arranged in an irregular sequence. Care was taken to have the same number of "*W* vs. *B*" as "*b* vs. *w*" combinations and to have each combination appear an equal number of times on the right and on the left. These latter sheets were used when "*b* vs. *w*" and "*W* vs. *B*" were being given in random order and, with slight modifications, could also be used for recording random presentation of "*B* vs. *b*", and "*w* vs. *W*"; of "*b* vs. *w*", "*W* vs. *B*", and "*B* vs. *b*", and finally of "*b* vs. *w*", "*W* vs. *B*", "*B* vs. *b*", and "*w* vs. *W*". In the presentation of the three and of the four discrimination-pairs in random order, each pair was presented the same number of times, and an equal number of times to the right and to the left. No pair was given more than three times in succession. The blank record-sheets were used when discriminations were being given in straight order and were headed with the stimulus-combination being presented.

The discrimination-time in tenths of a second was listed under the "*Time*" column after each response, a plus or minus for a correct or an incorrect response under "*Score*," the number of times that the animal turned its head in looking from one stimulus-patch to the other under "*Comparisons*," and observations on any special conditions obtaining, or unusual behavior of the animal, under "*Remarks*."

At the end of each day the percentage of correct responses for each subject at each experimental session was computed and recorded on a permanent record sheet, together with average time of response and percentage of response with comparison-behavior. These condensed score sheets are summarized in Tables 2-5.

TABLE 1

SHOWING ORDER IN WHICH THE VARIOUS STAGES OF THE PROBLEM WERE PRESENTED

	Series I				
Step 1	Step 2	Step 3	Series II	Series III	Series IV
b vs. *w*	*W* vs. *B*	*b* vs. *w* *W* vs. *B* Random order	*B* vs. *b* *w* vs. *W* & Both pairs random	*W* vs. *B* *B* vs. *b* *w* vs. *W* Random order	*b* vs. *w* *W* vs. *B* *B* vs. *b* *w* vs. *W* Random order

TABLE 2-A

RESULTS ON SERIES I FOR THE TWO CEBUS MONKEYS AND THE GIBBON

Order of trials	Cebus No. 1			Cebus No. 2			Gibbon		
	% of responses correct	Ave. time (Sec.)	% Comparison trials	% of responses correct	Ave. time (Sec.)	% Comparison trials	% of responses correct	Ave. time (Sec.)	% Comparison trials
Step 1									
Group 1	52	2.78	28	50	2.30	21	48	1.66	0
2	51	4.00	40	59	3.66	48	56	1.35	5
3	58	4.47	45	85 (70)	3.26	75	67	2.65	18
4	60	3.83	30	100	3.92	95	65	2.72	8
5	65	3.63	28				65	2.24	5
6	68	3.11	51				63	2.07	4
7	64	2.09	15				71 (135)	1.06	9
8	62	1.95	27				100	1.80	7
9	69	2.16	14						
10	68 (77)	1.79	14						
11	100	1.65	55						
Step 2									
Group 1	41	2.57	52	51	5.00	67	54	1.11	8
2	64	2.02	27	75 (35)	3.80	50	54	.97	9
3	72 (60)	1.62	50	100	3.45	87	67 (60)	1.19	24
4	100	1.65	55				100	1,29	81
Reinstatement Step 1				68 (220)	4.40	70	60 (480)	1.31	31
				100	3.85	80	100	.98	20

Each group (column 1) consists regularly of 100 consecutive trials except in cases in which the number of trials is indicated in parentheses in the second column. Also the last group on each step includes only the forty trials representing the norm.

TABLE 2-A (*continued*)

Order of trials	Cebus No. 1 % of responses correct	Cebus No. 1 Ave. time (Sec.)	Cebus No. 1 % Comparison trials	Cebus No. 2 % of responses correct	Cebus No. 2 Ave. time (Sec.)	Cebus No. 2 % Comparison trials	Gibbon % of responses correct	Gibbon Ave. time (Sec.)	Gibbon % Comparison trials
Step 3 (Split-series)									
Group 1	63	1.86	65	74	3.43	67	65	1.11	47
2	66	1.48	48				64	1.11	40
3	66	1.61	55				70	.95	48
4	76	2.30	75				55	.94	25
5	78	2.23	72				65	.94	47
6							65	1.14	53
(*Random S-Sets*)									
Group 1	75	2.09	57	86	2.95	81	73	1.12	69
2	85	2.02	74	94	2.77	85	80	1.53	87
3	84 (51)	1.92	92	100	1.70	78	86	1.63	96
4	100	1.67	85				100	1.30	80

TABLE 2-B

RESULTS ON SERIES I FOR THE THREE RHESUS MONKEYS

Order of trials	Rhesus No. 1 % of responses correct	Ave. time (Sec.)	% Comparison trials	Rhesus No. 2 % of responses correct	Ave. time (Sec.)	% Comparison trials	Rhesus No. 3 % of responses correct	Ave. time (Sec.)	% Comparison trials
Step 1									
Group 1	44	2.14	5	57	3.45	3	67	4.26	4
2	51	2.19	5	57	3.82	41	69	4.34	8
3	51	2.06	8	66	4.09	40	67	2.78	0
4	54	1.87	4	69	2.50	33	59	1.43	1
5	69	1.80	6	77	2.45	25	72	1.96	14
6	73	1.52	2	85	2.18	28	86	2.00	29
7	68	1.80	7	83 (90)	2.07	36	94	1.46	33
8	84 (50)	1.73	16	100	2.00	60	100	1.34	18
9	100	1.97	20						
Step 2									
Group 1	39	2.20	23	38	2.13	87	29	1.52	33
2	55	2.05	11	51	1.75	52	58	1.40	42
3	43	1.68	11	61	1.77	67	64	1.37	71
4	63	1.12	9	68	1.81	71	78	1.29	82
5	61	1.02	21	66	1.51	90	84	1.29	71
6	74 (120)	.96	31	74	1.23	83	82	1.15	55
7	100	.97	58	87	1.06	76	86	1.30	56
8				90 (70)	1.20	67	100	1.06	65
9				100	1.00	67			

Each group (column 1) consists regularly of 1 00 consecutive trials except in cases in which the number of trials is indicated in parentheses in the second column. Also the last group on each step includes only the forty trials representing the norm.

TABLE 2-B (*continued*)

Order of trials	Rhesus No. 1 % of responses correct	Ave. time (Sec.)	% Comparison trials	Rhesus No. 2 % of responses correct	Ave. time (Sec.)	% Comparison trials	Rhesus No. 3 % of responses correct	Ave. time (Sec.)	% Comparison trials
Step 3 (Split-series)									
Group 1	45	1.06	53	37	1.20	53	41	1.24	61
2	66	1.08	62	67	1.58	84	60	1.34	61
3	75	1.10	62	70	1.36	85	70	1.25	58
4	70	1.08	66	75	1.40	94	65	1.52	66
5	66	1.66	67	61	1.28	95	65	1.84	73
6	60	1.50	47	58	1.63	90	68	1.71	54
7				57	1.56	88	62	2.25	47
(*Random S-Sets*)									
Group 1	60	1.40	29	62	1.40	92	50	1.39	60
2	55	1.06	17	53	1.54	99	48	1.29	43
3	46	1.00	16	64	1.05	95	48	1.00	17
4	53 (120)	.98	26	59 (60)	1.08	98			
Pulling-in method (Split-series)									
Group 1	26	1.09	5	37	.97	17	35	1.61	18
2	44	1.12	32	39	.79	12	43	1.11	26
3	49	1.09	45	48	.85	28	55	.93	39
4	57	1.06	31	47	1.03	24	42	1.11	3
5	50	1.04	31	50	.85	25	44	.90	6

TABLE 3
RESULTS ON SERIES II FOR THE CEBUS MONKEYS AND THE GIBBON

Order of trials	Cebus No. 1 % of responses correct	Cebus No. 1 Ave. time (Sec.)	Cebus No. 1 % Comparison trials	Cebus No. 2 % of responses correct	Cebus No. 2 Ave. time (Sec.)	Cebus No. 2 % Comparison trials	Gibbon % of responses correct	Gibbon Ave. time (Sec.)	Gibbon % Comparison trials
B vs. b									
Group 1	39	1.77	74	42	1.47	85	54	1.48	73
2							38	1.51	75
3							36	1.97	86
4							58 (26)	1.80	89
Split-series									
Group 1	51	2.42	75	46	1.39	71	45	1.26	77
2	47	2.23	74	52	1.59	67	56	1.73	77
3	55	2.63	79	73	1.12	60	55	1.93	75
4	59	1.77	75	59	.98	55	52	1.35	62
5							48 (25)	1.30	64
w vs. W									
Group 1	61	1.75	82	56	.87	30	42	1.17	63
2	65	2.32	80	58	.85	62	45	1.40	51
3							53 (90)	1.35	59
Split-series									
Group 1	80 (44)	1.80	75	55	1.05	62	52	1.45	66
2				64	1.12	40			
3				58	1.32	67			
4				62	1.11	63			
5				71	.98	61			
6				49	.90	71			

Each group (Column 1) consists regularly of 100 consecutive trials except in cases in which the number of trials is indicated in parentheses in the second column. Also the last group of the series includes only the forty trials representing the norm. The *Split-Series* involved giving the two stimulus-combinations (*B* vs. *b* and *w* vs. *W*) in the same experimental period.

TABLE 3 (*continued*)

Order of trials	Cebus No. 1 % of responses correct	Ave. time (Sec.)	% Comparison trials	Cebus No. 2 % of responses correct	Ave. time (Sec.)	% Comparison trials	Gibbon % of responses correct	Ave. time (Sec.)	% Comparison trials
Reinstatement *w vs. W*									
Group 1				54	.85	61			
2				44	1.04	79			
3				47 (70)	.98	70			
Random S-Sets									
Group 1	61 (70)	2.25	83						
2	100	2.17	87						
Pulling-in *B vs. b*									
Group 1				45	1.71	39	43	1.51	63
2				49	1.12	55	52	2.39	79
3				84 (75)	1.04	86	46 (50)	2.87	94
4				100 (40)	1.05	100			
w vs. W				82 (60)	1.31	100			
Split-series									
Group 1							61	2.70	87
2							42 (120)	2.00	57
Random S-Sets									
Group 1				77 (30)	1.24	100	56	1.42	80
2				100	1.24	100			
Split-Series									
Group 1							49	1.53	63
2							54	1.60	65
Random S-Sets									
Group 1							45 (60)	1.95	65
Split-Series									
Group 1							48	1.83	70
2							39	1.77	60
3							42	2.14	54

TABLE 4

RESULTS ON SERIES III FOR THE TWO CEBUS MONKEYS

Order of trials	Cebus No. 1 Reaching technique % of responses correct	Ave. time (Sec.)	% Comparison trials	Cebus No. 2 Pulling-in technique % of responses correct	Ave. time (Sec.)	% Comparison trials
Mixed presentation						
Group 1	67	1.79	87	42	1.71	97
2	43	1.90	86	37	1.67	100
3	71	1.63	87	49	1.55	92
4				60	1.37	92
5				47 (55)	1.33	100
Random presentation						
Group 1	61	1.78	90	59	1.51	99
2	100	1.45	95	69	1.79	96
3				71 (42)	1.36	95
4				100	1.84	97

Each group consists regularly of 100 consecutive trials except in cases in which the number of trials is indicated in parentheses in the second column. Also the last group on the series includes only the forty trials representing the norm.

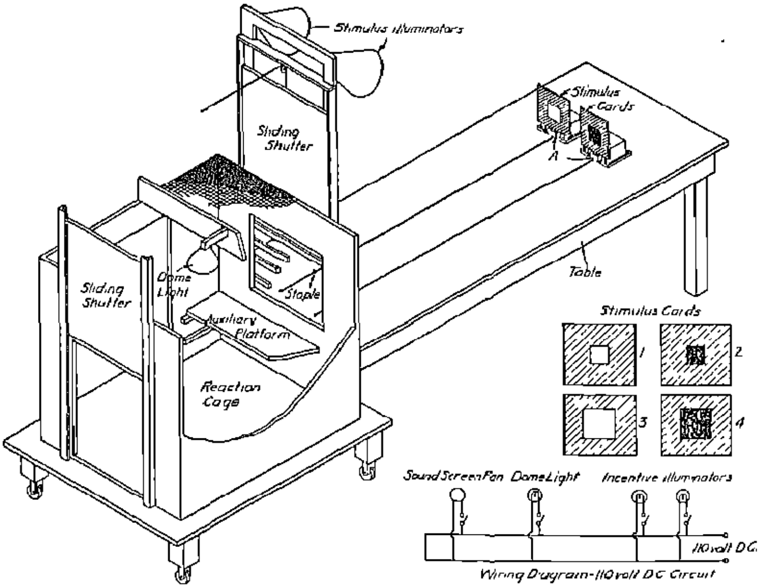


FIGURE 2

TABLE 5

RESULTS ON SERIES IV FOR THE TWO CEBUS MONKEYS

	Cebus No. 1 Reaching technique			Cebus No. 2 Pulling-in technique		
Order of trials	% of responses correct	Ave. time (Sec.)	% Comparison trials	% of responses correct	Ave. time (Sec.)	% Comparison trials
Mixed presentation						
Group 1	49	1.68	67	48	1.52	96
2	67	1.50	85	56	1.76	99
3	70	1.39	81	71	1.17	94
4	63 (57)	1.27	86			
Random Presentation						
Group 1	63	1.46	78	46	1.19	79
2	61	1.28	85	42	2.78	86
Split-Series	66 (85)	1.39	86	55 (80)	2.10	90
Mixed presentation						
Group 1	58	1.36	87	61	4.72	97
2	52	1.30	77	57	4.94	91
3	59 (80)	1.35	74			
Random presentation						
Group 1	51	1.30	80	70	2.75	96
2	58	1.31	85	64	2.07	93
3	62	1.35	71	48 (120)	3.75	90
4	60 (110)	1.40	73			

Each group consists regularly of 100 consecutive trials except in cases in which the number of trials is indicated in parentheses in the second column.

The order and nature of the pairs of stimuli used are shown in Table 1 and the several combinations of stimulus-factors, from series to series, are indicated graphically in Figure 3. It will be seen that all animals were started on discrimination "*b* vs. *w*". As soon as an animal reached the criterion of 40/40 correct choices, it was transferred to Step 2 of the first series. On mastering this, it was progressed to Step 3. If an animal was successful with the three steps of Series I, it was presented with Series II. If it learned this series it was given Series III and on meeting the criterion here, was finally presented with Series IV.

As all of the animals except two had great difficulty with the

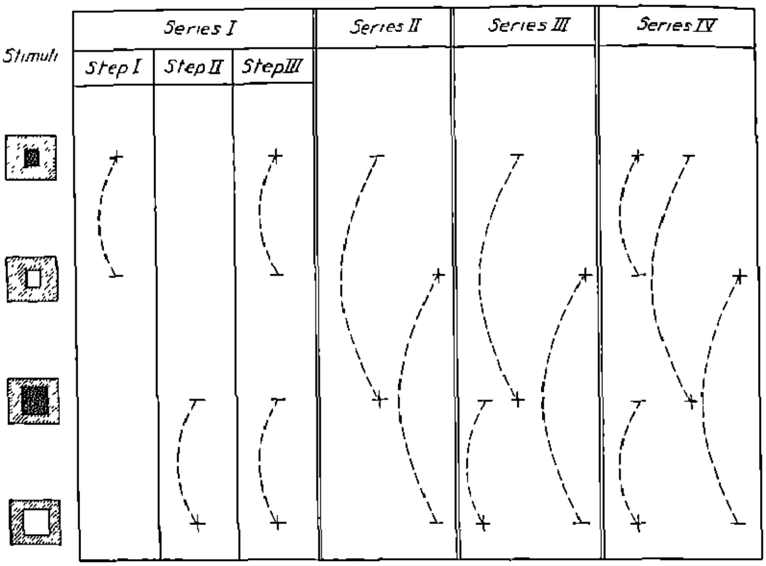


FIGURE 3

initial discrimination ("*b* vs. *w*"), the simpler discrimination of "no light vs. light" was presented. This stimulus-combination consisted merely of one side of the stimulus-incentive box being dark and the other side being lighted. All the animals quickly learned to reach into the dark side. After this had been learned, the "*b* vs. *w*" discrimination was resumed.

If at any point in the learning process an animal failed to make significant progress on a given series in 1000 trials, it was switched over to the pulling-in method. If in an additional 500 trials no progress was shown, the animal was considered to have reached the limit of its ability and experimentation on it was discontinued.

2. *Pulling-in Technique*

The animals were transferred from the reaching method to the pulling-in method whenever they seemed to reach their limit by the former method. Failure to show significant improvement in 1000 trials was taken as the point of transfer. The animal was continued for 500 trials on the same set of stimuli by the new method. (The gibbon was given 1000 trials instead of 500.) If it mastered the step it was continued on. As a matter of fact only one animal progressed significantly farther after being transferred.

The platform, food-trays, stimulus-blocks and cards, cord, and lighting used in this method have already been described. Only one of the food-trays was baited in a given trial, and if the animal first pulled the incorrect string it was not allowed then to pull in the correct box (except in preliminary trials) but was trained to withdraw into the reaction-cage. On rare occasions where the animal had made many consecutive errors it was allowed to pull in the second tray so that it would see that one of the trays was baited.

The new technique was introduced because of the possible advantages enumerated below:

1. Nissen used the pulling-in procedure in his similar study on chimpanzees and in order to compare our results with his it was thought advisable to check up our results by this method to see if any further learning occurred.

2. The response required of the animal in this latter method—the pulling in of a string to which a light tray was attached—might be considered a more natural response to a monkey than reaching through a closed glass door, especially in view of the very excellent results Klüver (31) has secured through the use of this procedure.

3. When the stimuli are presented by the pulling-in method (see Figure 2) there is no partition between the two stimuli as there is in the reaching set-up. It was felt that this condition in which the stimuli were not separated might offer better conditions for discrimination.

4. If an animal made a mistake in the reaching-method it simply found the door locked, and within approximately 40 seconds another trial would be given it. The punishment, therefore, was not severe and it seemed possible that the animal failed to be more careful in its response just for this reason. In the pulling-in method, however, the animal was subjected to a heavier penalty when it made an error. It had to pull in a tray at least 50 cms. before it discovered its error, and, in addition, there was approximately twice the interval (80 seconds) between trials, due to the longer time required by the experimenter to make the necessary adjustments.

5. It seemed possible that the gibbon might have been penalized by having the stimuli so close (reaching method) since the most

important function of vision in these animals is in connection with their brachiating mode of travel where the effective distances are rather great.

3. *Results and Comparison of Two Methods*

The results obtained by the pulling-in method have been given in the summary Tables 2-5. For the three rhesus monkeys there was no further learning on being transferred to the pulling-in procedure. As a matter of fact their percentage of correct responses dropped considerably and did not, during the 500 trials given, reach the *percentage they were obtaining at transfer.* While the gibbon made noticeable improvement on the discrimination-step upon which it was working at the time of transfer, it did not master this step. Its increased proficiency might easily be accounted for merely on the basis of the 1000 additional trials given. Cebus No. 2 went forward rapidly in its learning on being transferred to the pulling-in method, whereas it had long been at a standstill in its performance on the reaching set-up. It, however, did not go further by the pulling-in technique than Cebus No. 1 had by the reaching technique. They both arrived at exactly the same level of performance by these two different methods.

Apparently, then, the possible advantages of the pulling-in technique did not seem to be effective. The results indicate that the difference between the two methods is much less than might be supposed. There seems to be more tendency for position habits to be set up by the pulling-in method, and these habits are more difficult to break than in the reaching method. The reaching method has a great advantage over the pulling-in method from the point of view of time consumed in giving a trial, and simplicity of operation for the experimenter.

E. EXPERIMENTAL CONTROLS

In building up discriminatory habits of the kind described there are four possible sources of error:

1. The animal may make its responses on the basis of some visually detectable difference in the stimulus-cards other than the size or brightness differences being intentionally employed. This possibility was eliminated by testing the animal with a duplicate card after it had made 100 per cent correct responses on any

stimulus-combination. The way in which the scores of an animal fell below chance in going from one stimulus-combination to another also indicated that it was choosing on the basis of the differential dimension being given (brightness or size, as the case might be).

2. The animal may make its choice on the basis of olfactory or visual cues of the food-reward. In the reaching method, the closed glass doors cut off any olfactory cues from the food and the food was definitely out of sight of the animal, as was shown by certain test trials in which the animal responded to the empty side of the box although the other side was baited. In any event, both sides were baited with food throughout the training procedure, so that no *differential* cue could have been obtained even though the animal could see the food-reward. In the pulling-in method, the possibility of a differential olfactory cue was obviated by smearing the inside of both trays with the food-reward before the experimental session. Sight of the food by the animal was quite precluded by always placing it against the front wall of the tray.

3. The animal may make its choice on the basis of visual cues given by the experimenter. This possibility seems most unlikely in the present experiment, as the experimental room was darkened and the wire-screening about the reaction-cage served as a one-way light-screen. Furthermore, the experimenter stood behind the animal while it was making a response and the animal hardly ever turned its head before or while responding.

4. The animal may make its choice on the basis of auditory cues from the movements of the experimenter or from the shifting of the stimuli between trials. To avoid this possibility the experimenter adopted a standard routine of constant movements, making false settings when no movement was necessary. In addition, there was an electric fan which served as a sound screen.

III. RESULTS

The experimental results were secured between November, 1937, and June, 1938. As a matter of convenience the data will be treated in the following order: (1) Series I, (2) Series II, (3) Series III, (4) Series IV, and (5) Supplementary Analysis. The discussion will follow the order in which the various stimulus sets were presented, as may be seen by referring to Table 1 and Figure 3.

The data taken for each of the series consisted of three types of scores: (*a*) error scores, (*b*) discrimination-time scores, and (*c*) comparison scores, i.e., the number of trials on which the animal was observed to turn its head and look from one stimulus-patch to the other before responding.

A. SERIES I

The results of Series I for the two cebus monkeys and the Gibbon will be found in Table 2*A* and for the three rhesus monkeys in Table 2*B*. The grouping of the animals in the two tables is purely a matter of convenience and the tables should be examined together. The error scores indicate the progress of the animal in learning each discrimination problem. These are regularly given in terms of percentage correct in consecutive groups of 100 responses each, in the second column of the two tables. Where the groups do not consist of 100 trials, the actual number of trials is shown in parentheses in this same column. The final group on each step includes only the 40 trials representing the norm of mastery. The average time in seconds for the responses in each group of trials is shown in Column 3 of Tables 2*A* and 2*B*. The percentage of trials in each group on which the animal was observed to compare the stimuli before responding is shown in Column 4 of these tables.

Series I was made up of three steps (Figure 3). The first step consisted of learning the discrimination small *b* vs. small *w*, the second step consisted of learning the discrimination large *W* vs. large *B*, while the third step consisted of learning to respond to these two stimulus-combinations correctly when they were presented in split-series order and in random sequence. Before progressing from one step to the next the animal was required to make 40 consecutive correct responses on each preceding step.

The stimulus-combinations were presented in two distinct orders in

Step 3. In the early training on this step, groups of 10 trials on each of the two discriminations were given in the same experimental period (Split-series). The stimulus-combination given first at each period was varied from day to day. Later on, the two stimulus-combinations were presented in an irregular or random sequence so that the same combination never occurred more than three times in succession (Random *S*-sets).

In the case of Cebus No. 2 and the Gibbon, Step 1 was reinstated after the learning of Step 2 and before beginning Step 3. This procedure was followed by Nissen in the case of some of his chimpanzees. The practice of reinstating Step 1 was not followed for the rhesus monkeys as they had taken so many trials on the first two discriminations that it was thought better to start the split-series presentation immediately. Apparently this reinstatement phase was not very important in passing from Step 2 to Step 3 as is indicated by the fact that Cebus No. 1 mastered the series without such reinstatement and that both cebus monkeys mastered Series II without the reinstatement of the first step of that series. In a sense the first stage of Step 3 (Split-series) provided for reinstatement of the discrimination of Step 1. In this stage series of 20 trials of each daily period of 40 were given on Step 1 and the other 20 trials on Step 2. Thus provision was made through a long series of trials for further practice on the stimulus sets of Steps 1 and 2 before the random series were introduced.

It will be seen by reference to Tables 2*A* and 2*B* that both types of monkey and the Gibbon learned the first step of Series I. It will be noted that the percentage of correct responses for all the animals except one (Rhesus No. 3) was approximately 50 per cent on the initial group of Step 1, and that this percentage increased until the animals met the norm of mastery. There was likewise a tendency for the average time involved in making a discrimination to decrease as learning progressed, although this trend is not always consistent.

The comparison score represents a type of behavior similar to that reported by Crawford (8) involving noticeable movements of the head in comparing the two stimuli before reacting. Since only noticeable movements were recorded the scores probably represent the minimum of this type of behavior. It is possible, for example, that comparisons were made by eye-movements or by very slight movements of the head which escaped observation. As will be noted from the tables, there was a general tendency for the percentage of correct responses to increase as this type of behavior increased. The cebus monkeys began noticeably comparing the stimuli at a very much earlier stage in the learning of Step 1 than did the rhesus monkeys and the gibbon. In the first group of 100 trials, comparison-behavior occurred on an average of 24 trials for the cebus monkeys and on an average of only 3 trials for the rhesus monkeys and the gibbon. The cebus monkeys not only compared earlier in the training than did the other animals, but this behavior occurred on a much larger percentage of trials throughout the learning of Step 1.

Although all of the animals learned Step 1 there are considerable differences in regard to the number of trials required to reach perfect performance. The differences here do not follow species lines and hence should probably be regarded as individual differences in learning the test routine.

As will be noted in Tables 2*A* and 2*B* all of the animals were able to complete Step 2. Here the number of trials required to master the step do seem to follow species lines. The two cebus monkeys learned the discrimination in approximately 200 trials, the gibbon in 260 trials while the rhesus monkeys required from 650 to 750 trials.

The stimulus-combination used in Step 2 was somewhat antagonistic to that previously responded to in Step 1. This fact would naturally be expected to cause more or less interference in the learning of Step 2. While interference is not clearly indicated in comparing the total trials for learning the two steps, evidences of it did show up in the experiment. For the cebus monkeys, this interference is more or less masked by the pooling of trials in groups of 100 as in Tables 2*A* and 2*B* since the interference-effect was not of that long duration. However, if we consider the first 20 trials of each of the animals on the second discrimination we find that, instead of making a chance score of 50 per cent correct, Cebus No. 1 made a score of only 5 per cent and Cebus No. 2 of 25 per cent. The interference seemed to last for about 30 to 40 trials. The scores of the gibbon did not indicate this interference, but the effect was very marked for the three rhesus monkeys. Furthermore it lasted longer for the latter than for the cebus monkeys, so that it is distinctly shown in the data even when pooled in groups of 100 trials. By looking at Table 2*B* we find that these animals, instead of making chance scores or better in the first 100 trials on the second discrimination, made percentages correct ranging between 29 per cent and 39 per cent. The interference-effect thus persisted long enough to lower the average for the first complete group of 100 trials.

The same relations obtained in Step 2 as enumerated above in Step 1 as to the times scores and the comparison scores.

As will be noted there were two orders of presenting the stimuli in Step 3. In the first stage of the step they were presented in split-series order. As the cebus monkeys and the gibbon came to respond from 65 per cent to 75 per cent correctly, the mode of presentation was switched to a random order (Random *S*-sets). The rhesus monkeys reached a similar point of performance after several hundred trials on the split-series order. However, as this training was continued the percentage of correct responses decreased to around 60 per cent. Since these animals were getting worse instead of better they were transferred at this point to the random order presentation. The training of the rhesus monkeys on the split-series presentation was continued longer than for the cebus monkeys and the gibbon in order to give them every opportunity to master this simpler order before proceeding to the more complex random order presentation.

It will be seen from Table 2*B* that the three rhesus monkeys showed a very marked interference-effect on passing from Step 2 to the first stage of Step 3. Although the animals had all learned the two discriminations given in Steps 1 and 2 perfectly, when these same discriminations were given in the same experimental period, the percentage of correct responses in the first 100 trials fell below chance. They ranged in the three animals from 37 per cent correct to 45 per cent correct. A similar interference-effect was shown in the case of the cebus monkeys and the gibbon in the sense that their scores on the two discriminations when given together fell well below the 100 per cent mark. The interference was not as marked in these latter animals, however, as it was in the rhesus monkeys.

The two cebus monkeys and the gibbon mastered the second stage of Step 3 (Random *S*-sets presentation) up to the norm of 40 perfect responses in succession. The rhesus monkeys failed on this stage although they were given an equal amount of training. As a matter of fact during the last 40 trials of presentation to the rhesus monkeys, the percentage of their correct responses ranged only from 55 per cent to 60 per cent. This showed quite a marked deterioration from their best scores on the random order presentation which were attained soon after its introduction. These scores were 67 per cent correct in a series of 40 trials for Rhesus No. 1, 72 per cent correct for Rhesus No. 2, and 60 per cent correct for Rhesus No. 3. The fact that these higher averages were not maintained through the later training is one indication that the problem was too difficult for the animals. At any rate they were unable to form a stable response to the random presentation.

As the rhesus monkeys had done so poorly on the random presentation of the stimulus-combinations, it was discontinued when the animals were changed to the pulling-in technique, and the stimulus-combinations were again presented in groups (Split-series) as in the first stage of Step 3. It was our plan to introduce later the random order presentation by this new technique. However, since the scores made on the split-series were practically chance, the more difficult random series were omitted for the pulling-in schedule.

If we consider the number of trials required to master Step 3 as a whole (considering the reinstatement of Step 1 as part of that mastery in the case of Cebus No. 2 and the Gibbon) some interesting comparisons of the primate species being studied may be pointed out. Although the Gibbon finally mastered the step, it required about twice as many trials as did the cebus monkeys and stayed at the same level of performance for approximately 600 trials. Its difficulty with this step is in line with subsequent results which indicated that it was not able to go farther in learning the complex sets of discriminations. The failure of the rhesus monkeys on Step 3 is especially noteworthy in view of the fact that they were given twice as many trials as were required by the cebus monkeys to bring them to the norm of mastery. The failure of the animals at this point was doubtless due to the fact that this step is very much more complicated than Steps 1 and 2 (simple discrimination) and the split-series stage of Step 3.

B. SERIES II

The three rhesus monkeys were not carried on to Series II since each had failed on Step 3 of Series I. The two cebus monkeys and the gibbon were continued on this latter series and the results are summarized in Table 3. The task required here was the learning of two new discriminations (large *B* vs. small *b* and small *w* vs. large *W*) followed by response to these in random order (Figure 3). Series II, then, was very similar to Series I in that it required the learning of two new discriminations thoroughly isolated from one another. In the present case, however, the two sets of stimuli were not presented in steps as distinct as in Series I. The three steps in Series II were given more or less in conjunction with one another in accordance with Dr. Nissen's procedure with the chimpanzees at this point.

The first column of Table 3 gives the order of presentation of the various stimulus-combinations. The first step involved the presentation of *B* vs. *b* for a time. This was followed by a split-series order in which *B* vs. *b* and *w* vs. *W* were given for 400 to 500 trials. In the split-series arrangement each session of 20 trials consisted of 10 trials on each stimulus set. Usually two sessions a day were given. Following the split-series, training on *w* vs. *W* alone was given for several groups of trials. As indicated in the Order of Trials (Column 1), training on each of the two stimulus-combinations alone and in split-series was repeated as often as seemed necessary until it seemed feasible to present them in random order within the same session.

The amount of training on one or another stimulus-set varied from one animal to another according to the relative difficulty of each for the particular individual. It will be noted that Cebus No. 2 and the Gibbon received training by both the reaching and pulling-in techniques before tests with Random S-Sets were given. Shifting to the latter technique was not necessary for Cebus No. 1 since it had already learned the discriminations.

It will be seen from the results in Table 3 that the two cebus monkeys mastered Series II while the gibbon failed to do so. Cebus No. 2 required over twice as much training on this problem as did Cebus No. 1 and was not successful until shifted to the pulling-in technique. The Gibbon failed even though it was shifted to the latter method and was given in all several hundred more trials than Cebus No. 2. Not only did it fail to master the presentation of the two stimulus-combinations in random order but it was also unsuccessful on each of the component stimulus-complexes. This is true in spite of the fact that the Gibbon was changed back and forth from one stimulus-set to another in order to afford every opportunity for learning. The animal was given approximately 900 trials on the discriminations singly, 1200 trials on the split-series arrangement, and 200 trials on random presentation. The total of 2300 trials was considerably greater for the Gibbon than for either of the cebus monkeys. The amount of training given by the two techniques (reaching and pulling-in) was roughly the same. The Gibbon's best score on a random presentation of 40 trials (72 per cent) was made some 500 trials before the end of training. This score was not maintained thereafter and soon fell to below the chance level. At the termination of the experiment the animal was making less than a chance score on each of the two component stimulus-sets presented in the split-series arrangement.

The question naturally arises as to why Series II should be so much more difficult for the Gibbon than Series I, since both involved two stimulus-combinations in various presentational orders. It seems likely that the greater difficulty was due to two factors. In the first place, the initial discrimination in Series I was learned to a high degree of mastery before passing to the second one. In the present series, on the other hand, the second discrimination was introduced when the first one had been only partially learned and this would naturally prove confusing. In the second place, the emphasis on antagonistic cues, which would favor interference, was much greater in Series II. In fact, in the present case, the stimulus sets involved a combination of antagonistic cues in two dimensions—size and brightness. For example, the combination "small-black," which was correct in Series I became incorrect in Series II, and the combination "large-White" which was correct in the former series became incorrect in the latter. Series II involved, therefore, the reversal of response to two brightness-size combinations which had been previously learned. The fact that the original cue-combinations had involved long training and had been learned to a high norm of perfection (40 consecutive correct responses) doubtless favored a high degree of interference.

The importance of this interference factor in passing to Series II will appear from an analysis of Table 3. The amount of interference was greater for the Gibbon than for the cebus monkeys but was evident for all animals. For example, when the first new cue-combination (*B* vs. *b*) was presented, the Gibbon did not reach chance scores until over 300 trials had been given. This new combination was antagonistic in size-cue to the first stimulus-combination of Series I (*b* vs. *w*) and in brightness-cue to the second combination of the former series (*W* vs. *B*). Interference in connection with the second combination of Series II (*w* vs. *W*) cannot be clearly shown since this combination occurred first along with the other (*B* vs. *b*) in split-series presentation.

C. Series III

Since the Gibbon had failed on Series II only the two cebus monkeys could be carried forward to Series III. The problem here required correct response to three stimulus-combinations when presented in random order. These combinations included the pair just learned in the preceding series (*B* vs. *b*, and *w* vs. *W*) and the second discrimination learned in Series I, *W* vs. *B* (Figure 3). The animals were continued in the present task on the training technique (reaching or pulling-in) by which they had mastered Series II. The results are summarized in Table 4.

During Series III two sessions of 30 or 40 trials were usually given each day. These sessions were separated by an interval of approximately two hours. The training consisted of two stages as indicated in the first column of Table 4. In the first stage (mixed presentation) the order during a single session was as follows: A group of either 10 or 20 trials was given on the stimulus-combination from Series I (*W* vs. *B*) in order to reinstate it. This was followed by 20 trials in which the pairs of stimuli learned in Series II were presented in random order. In the second stage of the training (random presentation) all three combinations were presented in random sequence.

As can be seen from Table 4, Cebus No. 1 mastered the problem presented in this series in 400 trials and Cebus No. 2 in 642 trials. Although Series III would seem to be much more difficult than Series II, it was thus mastered by each of the cebus monkeys in less than half the number of trials required on the former. While Series III did not involve the learning of any new discriminations, it did require the animals to respond selectively to three combinations involving antagonistic cues. It should be noted that in the present series it was necessary to give quite an extended number of trials in mixed presentation before passing to the final random order. This was due to the fact that when the previously learned combination from Series I (*W* vs. *B*) was first introduced, the response to the random presentation of the pair from Series II broke down. Obviously each of the three combinations had to become quite independent before they could be reacted to in random order.

D. SERIES IV

The training in Series IV as in Series III was given in two stages. In the first stage (mixed presentation) the order during a single session was as follows: A group of 10 or 20 trials was given on the combination from Series I (*b* vs. *w*) in order to reinstate it. This was followed by 20 trials in which the three stimulus sets responded to in Series III were presented in random order (Figure 3). In the second stage of the training (random presentation) all four combinations were presented in random sequence.

Both cebus monkeys failed to meet the norm of 40 successive correct responses on this series. This was true in spite of the fact that a second long series of mixed presentation was introduced after the animals had failed on the first group of random trials. In the first set of 200 random trials, the score for Cebus No. 1 averaged 62 per cent correct, and for Cebus No. 2, 44 per cent correct. After further training on mixed arrangement, a second set of random trials was given. Although this set was continued for 410 (Cebus No. 1) and 320 (Cebus No. 2) trials, the final scores were not much above chance.

It is true, however, that from time to time a fairly high percentage of correct responses would be made for short runs. For example, both animals made a score of 82 per cent correct out of 40 trials at one point in the learning. But in both cases these "best scores" decreased during the 200 additional trials given Cebus No. 1 and the 250 trials given Cebus No. 2. In the last 40 trials the proportion of correct responses had dropped to 57 per cent and 37 per cent respectively. Since the animals were making progressively lower scores, the experimentation was discontinued at this point.

A comparison of Tables 4 and 5 will show that the animals were given much more training on Series IV than on the preceding series. Cebus No. 1 was given more than three times as many trials and Cebus No. 2 almost twice as many. This difference is even more striking if we compare the amount of training for random presentation only for the two series. For in this reckoning Cebus No. 1

was given over six times as much and Cebus No. 2 over twice as much. It is apparent, then, that adding the fourth stimulus-combination greatly complicated the problem. This can be easily understood since the positive cue (size or brightness) in one stimulus pattern was now always reversed in some other pattern. Work was continued longer on Cebus No. 1 because it consistently maintained a higher level of responses than the other animal. It was thought, therefore, that prolonged training might result in solution of the problem. Since the monkeys failed, it seems fair to conclude that the task of isolating cues for the appropriate responses to four stimulus-combinations involving antagonistic brightness-size aspects is beyond the ability of the animals. It is possible, of course, that types of monkeys other than the cebus and rhesus might succeed. *It is also possible that other individuals of the types tested might* possibly make a better showing.

E. SUPPLEMENTARY ANALYSES

Up to the present point we have been interested mainly in determining the limits reached by individuals and species in the task as a whole. These limits are dependent largely on the error scores as related to the norms used. There remain several further lines of analysis that should throw some light on various aspects of the problem. These will be discussed briefly in the following paragraphs.

1. *Comparison Scores*

A record was taken of each discrimination-trial on which an animal was observed to turn his head as though looking from one stimulus-patch to the other before making his choice. For lack of a better term, we have designated this data as comparison scores. It should be emphasized, however, that the scores represent only the obvious cases where the animal looked from one stimulus to the other. It was possible for a comparison to be made by eye-movements or very fine head movements which would have escaped observation. Our scores would represent, therefore, the minimum number of trials on which such behavior occurred.

Our interest in these scores is to determine whether the presence of such behavior was an aid to the animal in making correct responses. The analysis for each animal bearing on this problem will be found in Table 6. As indicated in the fourth column of the table,

TABLE 6

Showing Relationship Between Comparison-Behavior and Correct Response

Animal	Total number of discriminations	Per cent of total responses on which comparison-behavior occurred	Per cent correct on non-comparison trials	Per cent correct on comparison trials
Cebus No. 1	4757	66	49	70
Cebus No. 2	4887	68	58	79
Gibbon	4806	61	47	70
Rhesus No. 1	2970	26	52	82
Rhesus No. 2	3100	60	47	73
Rhesus No. 3	2980	40	52	78

when comparison-behavior was *not* observed the responses were seldom more than 50 per cent correct. In fact, when the scores for all animals are pooled the average is 51 per cent. When comparison-behavior was observed, however, correct responses ranged from 70 per cent to 82 per cent. It appears, therefore, that this type of behavior favors correct discrimination and that it probably involves genuine comparison of the two stimuli. A check of the pooled time scores shows that the choice on comparison trials was slightly slower in the case of each of the animals. This difference ranges from less than one per cent in Cebus No. 1 to 30 per cent in the Gibbon, the average being 12 per cent. As will be seen from Table 6, the cebus monkeys on the whole exhibited comparison-behavior more than either the Gibbon or the rhesus monkeys. This is the same order as the rating of the animals in their proficiency in the experiment as a whole.

Some further facts may be noted regarding the development of comparison-behavior during the experiment. An inspection of Tables 2-4, covering the results for the several series, shows that there was a trend for this type of behavior to increase as training progressed. It did not appear during the first 45 trials in any of the animals and it occurred at a much earlier stage of training for both of the cebus monkeys than for the rhesus monkeys. For instance, in the first 100 trials on Series I, the cebus monkeys compared the stimuli 25 per cent of the time whereas the rhesus monkeys compared in only 5 per cent of the trials. The gibbon also compared very little in the early stages of training. The highest comparison-scores for the different types of animals in a group of 100 trials was 87 per cent for the rhesus monkeys, 98 per cent for the cebus monkeys and 96 per cent for the Gibbon. At the time experimentation was discontinued the corresponding scores were 21 per cent, 96 per cent and 54 per cent respectively. In going from one series to another the increase in the percentage of comparison-trials was not perfectly consistent since it was affected by changing from one step to another in a given series and by shifts in training technique.

2. *Critical Trials*

Some tests were made to determine if changing the absolute size of the stimulus-patches would interfere with the pattern of response, and to ascertain whether the animals made use of relative or absolute sensory cues in the several discriminations investigated. Following Köhler, Klüver, and others we have designated such tests given upon completion of different stages of the training as "critical" trials. In these critical trials the animal was rewarded for response to either of the two stimuli simultaneously presented. The tests were given only after successful performances on the random presentation of Series I and Series II. The analysis is limited, therefore, to results on the cebus monkeys and the Gibbon. Critical trials in relation to the two series enumerated above will be discussed in order.

The first set of tests was made for the purpose of determining if altering the absolute size of the stimulus-patches in Series I would disrupt the responses of the animals. At this point the animals were responding perfectly to the two pairs of training stimuli which were black (Pos. S.) and white patches 5 cm. sq. and black and white (Pos. S.) patches 9 cm. sq. The corresponding sizes for the critical stimuli were 6 cm. for the small squares and 8 cm. for the larger squares. The brightness-factor was not changed, but the size-factor was here reduced as the new stimulus-pairs were presented in random order. Nevertheless each of the three animals responded correctly in 16 out of 20 trials when presented with these stimulus-combinations in random sequence. The difference between the sizes of the two pairs of training stimuli was then further reduced. The small squares were increased to 6.5 cm. and the large squares were decreased to 7.5 cm. When these stimuli were presented the Gibbon scored 16 correct responses out of 20, while the scores of the two cebus monkeys fell to the chance level. Apparently, then, a fairly marked change (75 per cent decrease) in the size difference proved to be only slightly disturbing to the Gibbon but completely disrupted the response of the cebus monkeys. In view of the small number of animals it is impossible to say whether this difference is truly generic or not.

The second set of tests was given to the cebus monkeys after they had learned the two discriminations of Series II (*B* vs. *b*, and *w* vs. *W*) in random sequence. These may be regarded as transposition tests since in the *new stimulus-combinations the size of the* larger stimulus was now the same as that of the smaller stimulus in the earlier training pairs. Brightness relationships in all cases remained the same as in the training series. The first test involved a change in size in the stimulus-pair from Step 1 of Series II (*B* vs. *b*). *The larger square in the critical* combination was 5 cm. and the smaller square was 2.8 cm. When given this test, Cebus No. 1 responded in 15 out of 20 trials to the larger black square, even though this had been the negative stimulus in the training series. This indicates a marked dependence on relational cues. Cebus No. 2, on the other hand, selected the smaller black patch in each of the 20 trials. This suggests that the animal had in the training series learned to *avoid* the small black patch on the basis of absolute size cues. This interpretation is strengthened by the fact that on an immediate re-test, the response to the original training pair was found to be perfectly maintained, as was also true in the case of Cebus No. 1.

The second transposition tests involved a size change of the same order in the stimulus-combination from Step 2 of Series II (*w* vs. *W*). In order to overcome any possible disturbance from the test just described, about 20 trials were given on the training pair before the critical trials were begun. In the present test the larger square in the critical pair had been the positive stimulus in the training pair (transposition in the positive direction), whereas in the preceding test the larger square had been the negative stimulus (transposition in the negative direction). In 10 out of 20 critical trials Cebus No. 1 responded on the basis of relative cues instead of in 15 out of 20 trials as in the preceding test. Cebus No. 2, however, chose the larger square (absolute cue) in 18 out of 20 trials just as in the *first transposition test.* Apparently, then, both animals tended to make use of the same type of cue in this test as in the previous one.

This is interesting in view of the fact that transposition in the positive direction is supposed to emphasize absolute cues, since most attention is presumably given to the positive stimulus. In reviewing the literature on this topic Spence (47) reports studies by Kafka (30), Bierens de Haan (2), Köhler (34), Klüver (32), and Spence (46), which support this view.

Before giving the third test in the transposition series, a few trials were given on the original training sets from Series II in random presentation in order to reinstate these discriminations. Since Cebus No. 1 showed considerable disturbance in this series it was eliminated from the present tests in order not to interfere with the later stages of the main problem. The third test involved giving the pairs of critical stimuli of Tests 1 and 2 in random order. In 15 out of 20 responses to this presentation, Cebus No. 2 chose on the basis of relational cues. It will be recalled that this animal had responded to both pairs when given singly in the preceding tests on the basis of absolute cues. No suggestion can be found as to why this shift occurred in this test since the stimulus pairs were exactly the same as those used previously. Apparently some factor associated with random presentation tended to favor the use of relative cues.

Some further facts relating to the use of cues in the present experiment may be drawn from the general training data. In Series I the animals all showed a tendency to use relational cues in passing from the first pair (*b* vs. *w*) to the second pair (*W* vs. *B*). The same tendency was noted in passing from the first pair (*B* vs. *b*) to the second pair (*w* vs. *W*) in Series II. However there was a tendency to make use of absolute cues in passing from the last pair in Series I (*W* vs. *B*) to the first pair of Series II (*B* vs. *b*). These trends are indicated by the fact that the scores made immediately after the shifting are below chance. Nissen reports corresponding relationships at these points in his work on the chimpanzee.

On the whole it would seem that the problem of the importance of relative versus absolute cues in animal discrimination is one that has not as yet been thoroughly analyzed. It seems likely that any one of a number of factors may determine the type of cue used in a particular situation. The work of Warden and collaborators (57, 58) on the ringed dove showed clearly that the method of using pairs of stimuli in brightness and size discriminations favors the use of relative cues. It was found, for example, that when the negative stimulus varied in both directions from the positive in an irregular order, the use of absolute response was favored. It was also found that, for the doves, the use of absolute cues is more difficult and represents a higher perceptual process. The same conclusion has been reached recently by Spence in regard to primates. In the present study there is some indication that individual differences are also important in determining the type of cue used. Moreover the evidence is clear that specific stimulus factors in the previous training pair determines to some extent the nature of the effective cue to be used in response to a later pair. For example, passing from *b* vs. *w* to *W* vs. *B* and from *B* vs. *b* to *w* vs. *W* favored the relational cues, while passing from *W* vs. *B* to *B* vs. *b* favored the absolute cues. Until more systematic studies have been made on animal perception, general principles regarding this problem can hardly be formulated.

IV. DISCUSSION OF RESULTS

The primary aim of this investigation was to study the behavior of the several primate types on a progressively complex discrimination task involving ambivalent and antagonistic cue-combinations. By ambivalent is meant that the cue (size or brightness) had a positive or a negative value in two different series, depending upon the other stimulus of the pair. The antagonistic factor involved the reversal of one of the cues (size or brightness) in the two stimulus-pairs of the same series. As already noted, we were able to proceed far enough to determine the limits of capacity in the case of each type tested. It seems important, at this point, to consider in some detail this aspect of the experiment. For easy reference, the results bearing on the limits reached by the several types have been presented in summary form in Table 6. As may be seen, these limits varied greatly among the primate types tested. In the present discussion an attempt will be made to explain this fact, in so far as possible, in terms of structural and behavioral differences. As a matter of convenience the following comparisons will be treated in order: (*a*) rhesus and cebus monkeys, and (*b*) monkey and gibbon.

A. COMPARISON OF RHESUS AND CEBUS MONKEY

As noted in Table 7, the rhesus monkeys were unable to complete Series I whereas the cebus monkeys mastered all series up to and

TABLE 7

SHOWING THE LIMITS OF CAPACITY OF THE SEVERAL PRIMATE TYPES ON A PROGRESSIVELY DIFFICULT DISCRIMINATION TASK

Animal	Series I Step 1 *b* vs. *w*	Series I Step 2 *W'* vs. *B*	Series I Step 3 *b* vs. *w* *W'* vs. *B* (Random)	Series II *B* vs. *b* *w* vs. *W'* & Both pairs random	Series III *W'* vs. *B* *B* vs. *b* *w* vs. *W'* (Random)	Series IV *b* vs. *w* *W'* vs. *B* *B* vs. *b* *w* vs. *W'* (Random)
Cebus 1	+	+	+	+	+	—
Cebus 2	+	+	+	+*	+	—
Gibbon	+	+	+	—*		
Rhesus 1	+	+	—*			
Rhesus 2	+	+	—*			
Rhesus 3	+	+	—*			

An asterisk indicates that the animal was tested by both the reaching technique and the pulling-in technique on this step. After the latter technique was introduced it was used in all later series for a given animal.

including III. The rhesus monkeys were given many more trials on Series I than were the cebus monkeys, but at the conclusion of training were making only about chance scores. Apparently, then, this level of difficulty marks the limit for the rhesus monkey in this type of task. It is interesting to note that each of the three animals (2 females, 1 male) stopped at this limit. The same was true of the two cebus monkeys, the limit for them being Series III. These results indicate, therefore, that the cebus monkey is vastly superior to the rhesus in the type of abstraction here required. This finding is in harmony with the work of Klüver (31), and the more recent investigations of Warden, Koch and Fjeld from the Columbia Laboratory, as yet unpublished.

In seeking to interpret the superiority of the cebus monkey we may turn first to such evidence as comparative anatomy has to offer. The main factors here would seem to be the brain-body weight relationship and the degree of convolutional development. The data relating to brain-body weights have been reported for both types by Koch (33) and are as follows: (*a*) Rhesus monkey, brain weight 126 grams (Tilney); body weight 4310 grams (Schultz). (*b*) Cebus monkey, brain weight 69.5 grams (Leche); body weight 1290 grams (Leche). The brain-body ratios of the two types, based on these measurements, are as follows: rhesus, 1:34.2; cebus, 1:18.5. The ratio of brain to body weight is thus almost twice as great for the cebus as for the rhesus monkey, and hence the former might be expected to be superior.

It seems impossible to say whether or not the convolutional development of the cebus is more marked than that of the rhesus monkey because the brain of the former has not been systematically studied. Most comparative anatomists have taken the brain of the rhesus as the monkey-type in tracing the evolution of this organ in the primates. Until more careful work has been done on the brain of the cebus monkey, it will be impossible to say whether the evidence from convolutional development will support that from brain-body ratio.

Doubtless some of the difference in capacity of the two monkey types on the present problem may be attributed to temperamental traits. Several investigators, including Koch (33) and Klüver (31) have noted rather marked differences in the attitude of the two types in work situations. Koch notes that the reactions of the cebus are less nervous and somewhat more deliberate than those of the rhesus, in a complex problem-box situation. He also remarks that the cebus monkey seems to be more docile in temperament.

In general, similar differences in attitude were noted in the present study. The rhesus monkeys seemed to be less attentive to the stimuli and showed a tendency to be restless and active in the reaction-cage between discriminations. They were more nervous than the cebus monkey and more easily startled by incidental noises. One animal in particular paced the cage continually between trials throughout the experiment. They appeared to lose interest in the task on occasions and at such times tended to respond more or less at random. The cebus monkeys, on the other hand, were very attentive, not easily distracted from the task, seldom responded carelessly, and were likely to show "surprise" when shifted to a new stimulus-combination. During the interval between trials they usually remained quietly on the platform waiting for the door to be raised. It is interesting to note in this connection that the frequency of comparison-behavior corresponds with the limits reached by the several types. This kind of behavior, which favored accurate discrimination, was much more marked in the cebus than in the rhesus monkey, the gibbon falling in between these two types in frequency of comparison-behavior.

Another type of behavior appeared in both the cebus monkey and the gibbon but never in the rhesus monkey. This involved an attempt to hold up the sliding door of the reaction-cage with one hand, after making an incorrect choice, so as to obtain the reward from the other incentive compartment. This behavior persisted in spite of the fact that the weight of the door always prevented the animal from succeeding.

An interesting aspect of behavior which occurred only in the cebus monkey was that they sometimes continued to work in spite of the fact that they did not eat the incentive after obtaining it. They would either reach for the incentive but leave it in the tray, or pick it up and throw it on the floor of the reaction-cage. This response was especially marked in Cebus No. 1. It was likely to appear when 50 trials or more per day were being given and probably meant that the animal was securing more reward (peanut) than it wanted. This was indicated by the fact that neither animal attempted to pick up the peanuts and eat them later on during the task or on leaving the reaction-cage. The animals seemed to be well motivated on such trials and showed no falling off in the score of correct responses. Apparently the work itself had become a sufficiently motivating factor for the animal, and this in itself might be regarded as indicating a high level of intelligence. A similar type of behavior has been noted in another cebus monkey by Carpenter and Locke (7). Such behavior was not exhibited by either the gibbon or the rhesus monkey.

A striking bit of "neurotic" behavior in one of the cebus monkeys may be noted in passing. The severe phase occurred in connection with the difficult combinations in Series IV and lasted for approximately two weeks. This series involved the random presentation of three previously learned stimulus combinations (Series III) and the addition of a fourth. After continued failure on this problem, the response to the three combinations carried over from Series III completely broke down. At this point the animal often refused to work, went to the rear of the cage and crouched in front of the exit door with its back turned to the exposed stimuli. Repeatedly it pushed and manipulated the door as if trying to escape. This negativistic phase of behavior alternated with a display of temper when the animal would pull the food trays into the reaction-cage and try to tear them apart with its teeth and hands. It would also bite fiercely at the adjustable platform on which it stood and was otherwise extremely restless and excited. After being dropped from the present problem, the animal remained in the laboratory during the summer and for the past four months has been used on a problem in instrumentation. It still seems to be markedly more nervous and excitable than before the onset of the attack. Similar behavior in the dog and other sub-primates has, of course, been reported by Pavlov (43), and Anderson and Liddell (1), and in the monkey by Fjeld (10). In all cases such emotional upset seems to appear when an animal is pushed to or beyond the limit of its ability in one or another task. The occurrence of this behavior on Series IV would thus seem to support our conclusion that this stage of difficulty was above the limit of the cebus monkey.

B. Comparison of Monkey and Gibbon

As is shown in Table 7 the Gibbon was unable to complete Series II which required a response to the random presentation of the second two pairs of stimulus-combinations involving stimuli ambivalent to those in Series 1. It thus ranked higher than the rhesus but considerably lower than the cebus monkey. The animal failed on Series II in spite of the fact that it was given several hundred more trials than the cebus monkeys. This poor showing of the gibbon is somewhat surprising inasmuch as it represents an intermediate type between the monkeys and the higher of the great apes. It is the more surprising since Nissen found that the chimpanzee shows much more facility on this same task than the monkeys showed in our own work. He found, for example, that the chimpanzee was able to master Series IV in somewhat less than half the number of trials which resulted in failure for our animals. Since he did not carry the experiment further, the limits for the chimpanzee are not known. Nissen's results would lead us to expect, therefore, that gibbons, as the lowest of the great apes, would outrank either of the monkey types.

In attempting to interpret the poor showing of the gibbon we may turn first to data in regard to comparative structural development. Leche gives the average brain weight of the adult gibbon as 130 grams, the average body weight as 9,500 grams, and the brain-body ratio as 1:73. We have just seen that this ratio was 1:18.5 for the cebus monkey so that the latter has a much larger brain in proportion to body size than has the adult gibbon. However, Leche calls attention to the fact that the ratio is considerably larger for young animals than for mature ones. This is due to the fact that the body of the gibbon increases greatly in size as the animal matures whereas the size of the brain grows very little after adolescence. Our specimen was pre-adolescent being only about three or four years old and weighing approximately 4500 grams. Nevertheless even if this weight is taken, the brain-body ratio is still less in the gibbon than in the cebus monkey. As mentioned earlier, however, one should not attach too much significance to this difference in brain-body ratio as there is no clear cut relationship between brain-body ratio and intelligence for types as widely separated as monkey and gibbon.

The difference in brain-body ratio in favor of the cebus monkey over the gibbon would seem to be in agreement with certain facts relating to convolutional and fissural development. Authorities are in agreement that the brain of the gibbon is more deeply convoluted, and that the occipital lobe in particular is more highly developed than in the monkey. Nevertheless they point out that in certain respects the brain of the gibbon is less highly developed than the monkey brain. For example, some of the brain indices are not as high for the former as for the latter. Tilney gives the forebrain index of the rhesus monkey as 84 per cent and this index for the gibbon as only 82 per cent. This index for the cebus monkey is not given. He further goes on to say:

> . . . It is evident that the brain of the gibbon in many respects is little more advanced in its development than that of the baboon or macacus. . . . In respect to fissural development, the gibbon shows less advance than either the baboon or macacus. . . . Certainly, the superficial appearance of the Gibbon's cerebral hemisphere places between it and the more highly complex endbrain of the great anthropoid apes a wide interval. . . .

Yerkes also stresses that the gibbon in structural as well as in behavioral characteristics is much nearer to the monkey than it is to the higher ape. From the standpoint of brain structure, therefore, there is little or no reason to expect that the gibbon would be above the monkey level of intelligence. In fact, the structural evidence on the whole would seem to agree fairly well with the results here obtained.

The question might be raised at this point as to whether or not the Gibbon was handicapped in any way in the laboratory environment. Our observations indicate that this does not seem to be the case. The animal appeared well adjusted to the work situation, was apparently as well motivated as the cebus monkeys, and showed marked and continuous attention to the various stimulus-combinations. Contrary to the observation of Boutan (5) that his gibbon was very easily fatigued and made irritable by a few tests, our animal retained attentiveness better and was less disturbed by continued testing than either the rhesus or cebus monkeys. This is clearly shown in a special series of five settings, carried out after the conclusion of the experiments, in which the animal was worked continuously for an hour and a half (100 trials) on five successive days. During these prolonged periods of exertion the Gibbon gave no evidence of fatigue or emotional excitement, was quiet and attentive throughout, and was still responding normally at the close of the periods. The monkeys, on the other hand, could not be given more than 50 trials a session without appearing to become fatigued, and sometimes they refused to work before the close of even shorter periods.

The only handicap that was observed for the Gibbon was its difficulty at first in picking up the food from the incentive tray because of its long fingers and rudimentary thumb. However this clumsiness was overcome in the first few days of work. The animal adapted very quickly to the pulling-in technique, although it had not been a subject in the previous patterned string tests as had the monkeys. At the end of only 10 preliminary trials it was responding by pulling in the trays in from two to three seconds.

Attention should be called to rather characteristic behavior expressions of an emotional nature which occurred rather frequently in the Gibbon. These were very similar to the pattern of response that occur as temper tantrums in children and consisted of slashing about wildly with the arms, crying out, and biting. These outbursts occurred sometimes in the experimental apparatus when the animal failed to secure the food-reward on several successive trials, and sometimes when the animal was being put back into its living-cage. The outbursts lasted for only a few minutes and were in no sense comparable to the more or less permanent neurotic state of Cebus No. 2, described above. At all other times the Gibbon was gentle and affectionate.

An interesting episode occurred one day which seemed to indicate high intelligence in the Gibbon. A small iron bar was used by the experimenter to threaten the animal in order to drive it into its living-cage. The animal, however, still refused to enter the cage and as soon as the experimenter turned away, came over and seized the rod and carried it to the top of the living-cages where it was quite inaccessible.

On the whole our results are in general harmony with such structural factors as are now at hand. This evidence would seem to be especially clear as regards the rhesus and cebus monkeys but considerably less so in the case of the gibbon as indicated above. The observed behavior of the several types also appear to support our main findings relative to the order of limits established. The weight of evidence would indicate that this order is clear and definite for the two types of monkeys. The rank position of the gibbon would appear to be less certain. In the first place, very few studies have been made on the behavioral capacities of this animal, hence there is little background for interpreting our results. As noted in the introductory section, such tests as have been made indicate that the gibbon shows no superiority over the monkey. In the second place, we used only a single animal in our experiment and the rank position here indicated must be regarded as tentative until a number of other individuals have been tested.

It should be mentioned that all the animals on which the present limits have been secured were in the very early stage of adolescence. This would mean that they probably had not reached full mental maturity. It is possible, therefore, that the limits for mature adults might be somewhat higher than were secured on these young animals. On the other hand, the task of securing limits on fully adult animals would probably present difficulties because they would likely be less docile.

The emphasis throughout the discussion has been on the discriminative aspect of the problem. Such emphasis would seem to be quite natural in view of the fact that the discrimination method was utilized and the actual data obtained fell within this field. It seems fairly obvious, however, that the behavior called for under our special conditions also bears directly on the process of abstraction. The problem here set involved much more than the setting up of a number of discrimination habits in sequence. This is certainly true after the introduction of ambivalent cues in Series II (Figure 3). From this point forward the problem required the isolation of specific cues having different values in the several contexts. Moreover, this isolation had to be complete enough to allow the cues to operate when specific cue-context combinations appeared in random order. In the more complex arrangements here employed, successful response would seem to require a level of abstraction closely akin to the early stages of reasoning. This view of the matter is in general harmony with present-day conceptions regarding the developement of reasoning in the child. The isolation and recognition of specific cues in varied contexts is essential to the reasoning process. It is clear that such a function requires not only a perception of the proper cue at the moment, but also a memory for the corresponding contextual factors. This analysis would seem to apply with equal force to the more difficult stages of the present problem. It seems fair to conclude, therefore, that this study demonstrates a degree of abstraction that is the essential precondition of reasoning, if not rudimentary reasoning itself, at least in the cebus monkey.

V. SUMMARY AND CONCLUSIONS

Several primate types (three rhesus and two cebus monkeys, and one gibbon) were tested on a progressively complex discrimination task involving ambivalent and antagonistic cue-combinations. The monkeys were all in the very early stages of adolescence and the gibbon was somewhat pre-adolescent. The task included four series of stimulus-combinations representing as many stages of difficulty. The reaching-method was used throughout except that a shift was made to the pulling-in technique when failure occurred at any stage. The norm of mastery for each series was 40 correct trials in succession in random presentation. The criterion of failure on a step was 1000 trials by the reaching-method, in which no significant progress was made, usually followed by 500 trials by the pulling-in technique. The results of the experiment based upon random order presentation appear to warrant the following conclusions:

1. *The rhesus monkeys failed to master Series I consisting of* two pairs of stimuli involving antagonistic cue-combinations.

2. The gibbon was successful on Series I, but failed to master Series II consisting of two antagonistic stimulus-combinations that represented ambivalence with respect to the pairs in Series I.

3. The cebus monkeys were successful on Series I, II, and III, but failed on Series IV comprising four pairs of antagonistic and ambivalent stimuli, although they showed some progress in the earlier stages of this series.

4. The rank-order of primate types on this task (cebus monkey, gibbon, rhesus monkey) corresponds to the order in which the animals ranked in the frequency in which comparison-behavior was observed. This type of behavior favored correct response.

5. In the final series, the task was complicated enough to set up marked "neurotic" behavior in one of the cebus monkeys, a mild *phase of which still remained seven months later.*

6. An attempt has been made to explain the main findings regarding limits of ability in the several primate types in terms of structural and behavioral differences. A tentative interpretation of the results in terms of abstraction and related processes has been offered.

REFERENCES

1. Anderson, O. D., & Liddell, H. S. Observations on experimental neurosis in sheep. *Arch. Neur. & Psychiat.*, 1935, **34**, 330-354.
2. Bierens de Haan, J. A. Über Wahl nach relativen und absoluten Markmalen (Versuche an Affen und Bienen). *Zsch. f. vergl. Physiol.*, 1928, **7**, 462-487.
3. ———. Die Baukunst eines niederen Affen (Cebus Hypoleucus Humb.). *Tijdscr. Nederl. Dierk. Vereen.*, 1930, **2**, 23-27.
4. ———. Zahlbegriff und Handlungsrhythmus bei einem Affen. *Zool. Jb.*, 1935, **64**, 267-288.
5. Boutan, L. Les deux méthodes de l'enfant. *Act. Soc. Linn. Bordeaux*, 1914, **68**, 3-146.
6. Breslaw, B., Barrera, S. E., & Warden, C. J. The effect of removal of the post-central convolution of the macacus rhesus monkey upon the delayed response. *J. Comp. Psychol.*, 1934, **18**, 207-226.
7. Carpenter, C. R., & Locke, N. M. Notes on symbolic behavior in a cebus monkey (Capucinus appella). *J. Genet. Psychol.*, 1937, **51**, 267-278.
8. Crawford, M. P. Brightness discrimination in the rhesus monkey. *Genet. Psychol. Monog.*, 1935, **17**, 72-162.
9. Drescher, K., & Trendelenberg, W. Weiterer Beitrag zur Intelligensprüfung an Affen (einschliesslich Anthropoiden). *Ztsch. vergleich. Physiol.*, 1927, **5**, 613-642.
10. Fjeld, H. A. The limits of learning ability in rhesus monkeys. *Genet. Psychol. Monog.*, 1934, **15**, 369-537.
11. Foley, J. P., & Warden, C. J. The effect of practice on the delayed reaction in the rhesus monkey. *J. Genet. Psychol.*, 1934, **44**, 390-413.
12. Gallis, P. Les animaux savent-ils compter? (Note préliminaire). *Bull. Soc. Roy. des Sci. de Liége*, 1932, **1**, 82-84.
13. Gellermann, L. W. Chance order of alternating stimuli in visual discrimination experiments. *J. Genet. Psychol.*, 1933, **42**, 206-208.
14. ———. *The double alternation problem: I. The behavior of* monkeys in a double alternation temporal maze. *J. Genet. Psychol.*, 1931, **39**, 50-72.
15. ———. The double alternation problem: III. The behavior of monkeys in a double alternation box-apparatus. *J. Genet. Psychol.*, 1931, **39**, 359-392.
16. ———. Form discrimination in chimpanzees and two-year-old children: I. Form (triangularity) *per se*. *J. Genet. Psychol.*, 1933, **42**, 3-27.
17. ———. Form discrimination in chimpanzees and two-year-old children: II. Form versus background. *J. Genet. Psychol.*, 1933, **42**, 28-50.
18. Guillaume, P., & Myerson, I. Recherches sur l'usage de l'instrument chez les singes: I. Le problème du détour. *J. de Psychol.*, 1930, **27**, 177-236.
19. ———. Recherches sur l'usage de l'instrument chez les singes: II. L'intermédiaire lié à l'objet. *J. de Psychol.*, 1931, **28**, 481-555.
20. ———. *Recherches sur l'usage de l'instrument chez les singes:* III. L'intermédiaire indépendent de l'objet. *J. de Psychol.*, 1931, **31**, 497-554.

21. Haggerty, M. E. Imitation in monkeys. *J. Comp. Neur. & Psychol.*, 1909, **19**, 337-455.

22. Harlow, H. F., Uehling, H., & Maslow, A. H. Comparative behavior of primates: I. Delayed reaction tests on primates from the lemur to the orang-outan. *J. Comp. Psychol.*, 1932, **13**, 313-343.

23. Harlow, H. F. Comparative behavior of primates: III. Complicated delayed reaction tests on primates. *J. Comp. Psychol.*, 1932, **14**, 241-252.

24. Harlow, H. F., & Settlage, P. H. Comparative behavior of primates: VII. Capacity of monkeys to solve patterned string tests. *J. Comp. Psychol.*, 1934, **18**, 423-435.

25. Hartman, C. J. & Straus, W. L. Anatomy of Rhesus Monkeys. Baltimore: Williams & Wilkins, 1933. Pp. 383.

26. Jackson, T. A., Stonex, E., Lane, E., & Dominguez, K. Studies in the trans-position of learning by children: I. Relative vs. absolute response as a function of amount of training. *J. Exper. Psychol.*, 1938, **23**, 578-600.

27. Jacobsen, C. F. Functions of frontal associations area in primates. *Arch. Neur. & Psychiat.*, 1935, **33**, 558-569.

28. ———. Studies of cerebral function in primates: I. The functions of the frontal association areas in monkeys. *Comp. Psychol. Monog.*, 1936, **13**, No. 3. Pp. 60.

29. Johnson, H. M. Visual pattern discrimination in the vertebrates: II. Comparative visual acuity in dog, the monkey, and the chick. *J. Anim. Behav.*, 1914, **4**, 340-361.

30. Kafka, H. Beitrag zur psychologie eines niederen Affen: Grössunterscheidung bei cerococebus fuliginosus. *Zstch. f. vergl. Physiol.*, 1931, **15**, 71-120.

31. Klüver, H. Behavior Mechanisms in Monkeys. Chicago: Univ. Chicago Press, 1933. Pp. xvii + 387.

32. ———. The equivalence of stimuli in the behavior of monkeys. *J. Genet. Psychol.*, 1931, **39**, 3-27.

33. Koch, A. M. The limits of learning ability in cebus monkeys. *Genet. Psychol. Monog.*, 1935, **17**, 163-234.

34. Köhler, W. Aus der Anthropoidenstation auf Teneriffa: IV. Nachweis einfacher Strukturfunktionem beim Schimpansen und beim Haushuhn. Über eine neue Methode zur Untersuchung des bunten Farbensystems. *Abhn. Preuss. Akad. Wiss.*, 1918, 1-101.

35. Leche, W. Der Mensch. Jena: Gustav Fischer, 1922. Pp. 390.

36. Locke, N. M. Color constancy in the rhesus monkey and in man. *Arch. of Psychol.*, 1935, No. 193. Pp. 38.

37. Maslow, A. H., & Harlow, H. F. Comparative behavior of primates: II. Delayed reaction tests on primates at Bronx Park Zoo. *J. Comp. Psychol.*, 1932, **14**, 97-107.

38. Neet, C. C. Visual pattern discrimination in the macacus rhesus monkey. *J. Genet. Psychol.*, 1933, **43**, 163-196.

39. Nellman, H., & Trendelenburg, W. Ein Beitrag zur Intelligenzprüfund niederer Affen. *Ztsch. f. vergl. Physiol.*, 1926, **4**, 142-200.

40. Nissen, H. W. Concurrent antagonistic visual discrimination habits in the chimpanzee. *Psychol. Bull.*, 1933, **30**, 704.

41. ———. Equivalence and ambivalence of stimuli in chimpanzees. *Psychol. Bull.*, 1934, **31**, 617-618.

42. Ohtsuka, N. Über die absolute und relative Wahl beim Affen Cercopitheeus sp. *Acta Psychol. Keijo,* 1937, **3**, 32-44.

43. Pavlov, I. P. Lectures on Conditioned Reflexes: Twenty-five Years of Objective Study of the Higher Nervous Activity (behavior) of Animals. New York: International Publications.

44. Révész, G. Experimental study in abstraction in monkeys. *J. Comp. Psychol.*, 1925, **5**, 293-343.

45. Robinson, E. W. A preliminary experiment on abstraction in a monkey. *J. Comp. Psychol.*, 1933, **16**, 231-236.

46. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430-444.

47. ————. Experimental studies of learning and the higher mental processes in infra-human primates. *Psychol. Bull.*, 1937, **34**, 806-850.

48. Tilney, F. The Brain from Ape to Man. (2 vols.) New York: Hoeber, 1928. Pp. xvii + 473; xv + 474.

49. Tinklepaugh, O. L. Multiple delayed reaction with chimpanzees and monkeys. *J. Comp. Psychol.*, 1932, **13**, 207-243.

50. Tellier, M. Le macaque saisit-il le rapport logigue? *Bull. Soc. Roy. Sci. de Liége,* 1932, **1**, 227-231.

51. Verlaine, L. Le relatif et l'absolu chez le macaque: La grandeur moyenne. *Bull. Soc. Roy. Sci. de Liége,* 1935, **4**, 90-96.

52. ————. De la connaissance chez le macaque: La substance. *Bull. Soc. Roy. Sci. de Liége,* 1935, **4**, 239-242.

53. ————. Le caractère analytique de la perception chez le macaque. *Bull. Acad. Belg. Cl. des Sci.*, 1935, **21**, 744-758, 800-816.

54. ————. La vision de formes chez le macaque: Syncrétisme, analyse et synthese. *Mém. Acad. de Belgique Cl. des Sci.*, 1935, **14**, 1-85.

55. Verlaine, L., & Gallis, P. L'intelligence chez les singes inferieurs. *Mém. Soc. Roy. Sci. de Liége,* 1932, **17**, 1-48.

56. Warden, C. J., Jenkins, T. M., & Warner, L. H. Comparative Psychology: Vertebrates. New York: Ronald Press, 1936. Pp. x + 560.

57. Warden, C. J., & Rowby, J. B. The discrimination of absolute versus relative brightness in the ring dove, Turtur risorius. *J. Comp. Psychol.*, 1929, **9**, 317-337.

58. Warden, C. J., & Winslow, C. N. The discrimination of absolute versus relative size in the ring dove, Turtur risorius. *J. Genet. Psychol.*, 1931, **39**, 328-341.

59. Woollard, H. H. The differentiation of the retina in the primates. *Proc. Zool. Soc. London,* 1927, No. 1, 1-17.

60. Yerkes, R. M., & Yerkes, A. W. The Great Apes; A Study of Anthropoid Life. New Haven: Yale Univ. Press, 1929. Pp. xix + 652.

61. Yoshioka, J. G. Studies of abstraction with the chimpanzee. *Psychol. Bull.*, 1934, **31**, 601.

62. Yudin, H. C., & Harlow, H. F. Comparative behavior of primates: V. Delayed reactions in primates in horizontal and vertical planes. *J. Comp. Psychol.*, 1933, **16**, 143-147.

63. Zuckerman, S. Functional Affinities of Man, Monkeys and Apes. New York: Harcourt, Brace; London: Kegan Paul, 1933. Pp. 203.

64. ————. The Social Life of Monkeys and Apes. New York: Harcourt, Brace, 1932. Pp. 356.

$7.00 per volume
Single numbers $3.00

QUARTERLY

November, 1939
Volume 21, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota

Charlotte Bühler
Universität, Wien

William H. Burnham
Clark University

Cyril Burt
University of London

Leonard Carmichael
Tufts College

Ed. Claparède
Université de Genève

Edmund S. Conklin
Indiana University

Sante De Sanctis
R. Università di Roma

Arnold Gesell
Yale University

William Healy
Judge Baker Foundation,
Boston

Leta S. Hollingworth
Teachers College,
Columbia University

Buford Johnson
The Johns Hopkins
University

Harold E. Jones
University of California

Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal
College

K. S. Lashley
Harvard University

A. R. Luria
Medico-biological
Institute, Moscow

Toshio Nogami
Kyoto Imperial
University

Henri Piéron
Université de Paris

George D. Stoddard
State University of
Iowa

Calvin P. Stone
Stanford University

Lewis M. Terman
Stanford University

Godfrey Thomson
University of
Edinburgh

E. L. Thorndike
Teachers College,
Columbia University

C. J. Warden
Columbia University

John B. Watson
New York City

Helen Thompson
Woolley
Teachers College,
Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

Provincetown, Massachusetts
Copyright, 1939, by The Journal Press

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the Act of March 3, 1879.

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879.

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publications shall be immediate and in which other publication experiments may be carried out. Beginning with January, 1938, the following methods and terms will be in force:

(1) The subscription price will be $7 per volume, two volumes per year.

(2) The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3) Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4) The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5) Each author will be charged the actual manufacturing costs of any zinco or half tone cuts.

(6) The publication charges will be as follows:

4	page	signature	$16
8	"	"	20
12	"	"	30
16	"	"	40
20	"	"	50
24	"	"	60
28	"	"	70
32	"	"	80
36	"	"	90
40	"	"	100

These charges are tentative and will be adjusted as justified by the subscription list.

(7) The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $3.00

QUARTERLY

November, 1939
Volume 21, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not *be resold for a period of five years from date of publication.*

NOVEMBER, 1939

(Manuscripts are printed in the order of final acceptance)

THE SOCIAL-SEX DEVELOPMENT OF CHILDREN 461
By Elise Hatt Campbell

Copyright 1939, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under the act of March 3, 1879

Reentered as second-class matter April 15, 1937, at the post-office at Provincetown, Mass., under Act of March 3, 1879.

Published as a separate and in *Genetic Psychology Monographs*, 1939, **21**, 461-552.

THE SOCIAL-SEX DEVELOPMENT OF CHILDREN*

ELISE HATT CAMPBELL

The Merrill-Palmer School

Acknowledgments 463

I. Introduction 465

II. Summary of the literature 469

III. Method 487

IV. Reliability of the scale 509

V. Relation of social-sex development to other aspects of development 519

VI. Use of the scale 521

VII. Characteristics of social-sex development in various age groups, according to the scale 523

VIII. Sex differences and similarities in social-sex development . 527

IX. Sequence of certain aspects of social-sex development . . 529

X. Comparison of results with those of other studies . . . 533

XI. Case studies of social-sex development 535

XII. Summary and conclusions 545

References 549

---

*Received in the Editorial Office on July 27, 1939, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

ACKNOWLEDGMENTS

I am indebted to Isabelle Unruh, who aided in the groundwork of this study through her Master's thesis; to Joan Millington and Julia Dawson, who made some of the observations; to E. Virginia Van Dyne, who made the statistical correlations; to Dr. Paul H. Furfey and Dr. Wendell Vreeland, who suggested methods; to Dr. Katherine E. Roberts and others of the Merrill-Palmer staff who have read the completed study critically; and to Dorothy Tyler, who edited the study. These persons are not responsible, however, for any of the faults of the study. Much of the statistical work was accomplished through help provided by the National Youth Administration, and the author wishes to acknowledge the great assistance given by this governmental agency.

ELISE HATT CAMPBELL

THE MERRILL-PALMER SCHOOL
71 EAST FERRY AVENUE
DETROIT, MICHIGAN

I. INTRODUCTION

The adult who observes groups of children over a period of years and finds a pattern of behavior repeated in each group is inclined to see a native reaction pattern at work. Such an explanation occurred to the staff of the Merrill-Palmer recreational clubs as year after year they observed boys and girls reacting to each other at one stage with indifference, at another with hostility, and finally, at adolescence, with growing interest.

The recreational clubs were established in 1929 as a means of following up the development of children who had attended the Merrill-Palmer nursery school, and of permitting the observation and study of children of school age in connection with a program relatively as free as that of the nursery school. The program has been developed experimentally in accordance with the desires and suggestions of the children. The different age groups meet weekly, and thus boys and girls are brought together at each age from the kindergarten level through the high-school years.

The membership of each club was decided upon the basis of age ranges in which there would be homogeneous play interests, but the boys and girls in the group were allowed complete freedom to mingle or separate. Under these conditions, as one age group was replaced by an oncoming one, the staff began to notice that the youngest group, aged five to about seven, seemed to ignore sex in choosing play groups. Boys chose girls in their dramatics, or on their "sides," as freely as they chose boys. At eight years, however, the picture began to change, and by 10 or 11 boys and girls almost completely segregated themselves from each other. This stage of segregation began with a haughty aloofness, became apparent contempt and then active hostility, and then changed to a shy withdrawal which seemed to mark the end of this period and the beginning of the adolescent heterosexuality after puberty. This change occurred earlier in girls than boys but was easy to observe in both. It is a pattern familiar to everyone who has read the fiction and scientific writings describing adolescence. The primping of the girl and the cleanliness and care for personal appearance of the boy are some of the usual external characteristics.

As these changes were observed it occurred to the author that it might be interesting and practical to make more methodical observations and to analyze the separate features of this pattern. It was practical because the results might show whether there was a sound basis for our belief that recreational clubs for preadolescents should be coeducational, in contrast to the national programs set up on the assumption that there was good reason for not making them so. Were we contending against biological forces too strong for us and actually provoking hostility between the sexes by putting them together, or were we facilitating the change from the unisexual period to the heterosexual one by providing a means for it to take place with greater ease and less embarrassment and conflict? The writings of most authors on adolescence led us to believe that the latter was true.

There was also the theoretical interest in the nature of this pattern. Was it definitely tied up with the biological nature of the child? Were the physiological changes at puberty always preceded and followed by social reactions of the sexes to each other such as we saw repeatedly in our groups? Would children in other strata of American society and under different recreational conditions show the same behavior? And would children under entirely different cultural conditions display the same developmental picture? To answer these last two questions to any extent at all, reliance had to be placed on descriptions of childhood in other cultures, and especially on the works of Margaret Mead. Indeed, the works of this distinguished anthropologist were the strongest influence in opening our eyes to the possibility that social-sex development may be more culturally than biologically conditioned. It is perhaps proper as well to look for a "child culture" *per se* in the midst of each particular society.

To define the purpose of the present study more precisely: It attempts to analyze the development of the social relationships between boys and girls between the ages of 5 and 16 years, and, through this analysis, to arrive at a normative picture of the social pattern through which each individual develops into a heterosexual being, with individual variations. The data are limited to observable group patterns. There was no attempt to analyze the children by interview, through questionnaires, or other means of self-analysis. No inquiry was made as to the sex knowledge or attitudes of the children or their intimate friendships outside the clubs. The study is limited to the observable social relationships between normal American boys and girls in a free recreational situation. It is not directly *a study of sex differences, except when these occur in boys' and girls'* relationships to each other.

In a sense, the study may be considered one of the development of heterosexuality throughout later childhood and early adolescence.

So far as the author has been able to determine, the term "social-sex" has not been used in the published literature, but Pressey (52), in his excellent psychology of childhood, calls attention to the "sex-social" behavior and uses the term in precisely the same sense as the term "social-sex" is used here. Unruh (68), in an unpublished Master's thesis, defines social-sex development as "the ability of the individual to progress gradually through each phase of life, according to the social pattern of his environment, to a final heterosexual adjustment."

II. SUMMARY OF THE LITERATURE

A. OBJECTIVE STUDIES

Studies based on objective data describing social-sex patterns are rare. Furfey's publications on developmental age most nearly approach our concept. A Doctor's dissertation by Sullivan (64), also from the Catholic University of America, gives somewhat the same data for girls as Furfey's studies do for boys.

In *The Gang Age,* Furfey (21) describes the social-sex relations of the pre-adolescent boy—the cessation of play with girls and the sharp separation of the sexes typical of the period which begins between 8 and 11 years and ends at puberty. In *The Growing Boy,* he (22) has divided social-sex development into three stages. At six, boys mingle freely with girls; indeed, 88 per cent played with girls. Their relationship is essentially different from that in adolescence. Their interests are similar and they fight and play together without much consciousness of sex. Furfey believes that there may be a trace of eroticism in their relationships at this time, but it cannot be said to resemble that in adolescence. The boy at this time shows beginnings of masculine interests in that he begins to identify himself with his own sex, but he plays with girls on a basis of equality, usually games typical of the preschool age, such as dramatics. Furfey describes the boy of eight as still willing to play with girls in dramatic play; at least 73 per cent still play with girls. At 10, however, which marks the beginning of the gang age, and until puberty or about 14, there is a strong distaste for playing with girls. "The ultimate insult to the lad at this period is to call him a sissy and his idea of absolute zero in amusement is to call on a girl." The boy of 12 shows an increase in loyalty to his own sex and distaste for the opposite sex; only 20 per cent played with girls. At 14, boys show extremely heterogeneous patterns, owing probably to the variation in the age at which puberty occurs. A sufficient number had passed beyond the gang age to show 17 per cent who had had "love affairs." The boy of 16 has progressed much farther toward heterosexuality, yet he may still be hesitant and shy with girls; 36 per cent had had love affairs. These norms were derived from the administration of the scale for measuring developmental age to an unselected group of 450 grammar-school and high-school boys.

Sullivan's (64) results show that at six and eight girls often included boys in their play groups. At 10, 28 of the 35 girls preferred to play with other girls, two preferred the other sex, and five were neutral. At 12 the same proportions held. Two years later, at 14, 30 of the 35 girls showed a definite adolescent interest in boys, and at 16 interest in members of the opposite sex was well established.

In a group of gifted children, Burks and others (9) report an interest in the opposite sex, in boys, increasing from 25 per cent at 11 and 12, to 55 per cent at 13 and 14, and 88 per cent at 19 and 20. The comparative figures for girls were higher at almost all stages, ranging from 60 per cent at 11 and 12, to 78 per cent at 19 and 20. The earlier Stanford (65) study of gifted children has one table showing certain trends toward heterosexuality as the gifted and control children matured. The data show these trends in three age groups, and indicate that, perhaps owing to an earlier puberty, gifted children were somewhat ahead of control children in making heterosexual adjustments.

Though for various reasons the data from these several studies are not entirely comparable, it is evident that all agree in showing an increased heterosexuality between the ages of 12 and 16, with the girls ahead of the boys at each stage. It is fairly evident also that at 14 the girl is well on the way toward completing this adjustment, while the boy is still far behind.

Hildreth's (27) interesting study of the interests of 9th and 12th grade boys and girls from two groups, public and private schools, compares preferred activities and preferred associates at the two levels and in the two groups, and shows some evidence of earlier maturity among the private-school girls than among public-school girls. Though the data offer little basis for comparison with those of the studies already mentioned, they also show a progressive degree of heterosexuality, with the girl more advanced throughout adolescence.

Dimock's (14) careful analysis of the time schedules of 200 adolescents do not show, except by inference, the relative heterosexuality of the 12 and 16 year old. His view of heterosexual growth is that, while it is given some impetus by the physiological changes at puberty, it can be acquired only by means of the social *milieu* which fosters it. Dimock plans to publish further studies in this series.

Other studies throw some light on the problem. Koch (32) discovered a tendency toward preference for their own sex among 4-year-olds, though "a wholesome heterosexual interest was apparent." Green's (25) study of friendships among preschool children shows, even at this early age, a greater tendency toward the unisexual friendship. Parten (50) also found a strong tendency toward the unisexual friendship among preschool children. Hagman (26) found that younger preschool children showed no strong preference for either sex, but that later in this period a preference for like sex became apparent, especially in boys.

Lehman and Witty (33) found that at all ages, but especially between 6 and 12, boys tend to play with boys and girls with girls. Games tended to be sex-linked in middle childhood. "Dates" were not common among girls or boys in early adolescence, but, though still low on the list of boys' activities in late adolescence, jump to a high position with girls between the ages of 16 and 20. A marked rise in girls' preferences for dates occurs at 16, and two years earlier for social dancing. From unpublished material by Follenberger quoted by Shuttleworth (60), it appears that there is a definite trend toward increasing heterosexuality in boys from 9 to 18 years. Between 9 and 11, less than half the boys agreed that "girls are OK"; from 11 to 15, 50 to 60 per cent agreed that they were; and a marked increase occurred between 16 and 18. In a study of motivation, Maller (37) found that among children 2 to 14 years of age, work for one's own sex superseded work for oneself, one's team, one's school class, or an arbitrary team assigned by a teacher, as a motive for accomplishment, a result indicating the strength of unisexual interest during preadolescence. Murphy and Murphy refer to a Russian study by Cherkassova (12) showing that children of 13 to 16 evinced great hostility toward working as partners with the other sex in teams of two, while teams of two of the same sex were harmonious. The boys showed even greater hostility than the girls, which may be further evidence of the earlier heterosexuality of girls. Among the oldest children, the mixed teams tended to be more amicable, and arguments often ended in friendly interest. Another foreign study, by Fuxloch (23), mentioned in the literature but unobtainable to the author, showed that among children of school age a companion of the same sex was preferred. A study by Stone and Barker (61), reported in abstract in the proceedings of the American Psychological Association for 1936, appears to contain some data relevant to the social interests of pre- and post-menarcheal girls.

B. NON-OBJECTIVE WRITINGS (EXPRESSING POINTS OF VIEW ABOUT THE DEVELOPMENT OF SOCIAL-SEX RELATIONSHIPS)

It is difficult to find a child psychology book which does not mention the necessity for achieving heterosexuality in adolescence. Different reasons are ascribed for the existence of the universal social-sex pattern of childhood. Frankwood Williams (71) states strongly the need for achieving heterosexuality during adolescence; by this he means "a healthy, adult level of sexuality in which the primary sex interest of the individual is in the opposite sex." The preadolescent child, he states, is not heterosexual, and if the adolescent does not, through social contact, become interested in the opposite sex, this adjustment is not likely to be normal later. Pechstein and McGregor (51) describe the belief of Hall and his disciples that social-sex growth is saltatory in character, one phase of it being the withdrawal from the other sex just before puberty, ascribable to the earlier puberty of the girl. Pechstein and McGregor themselves believe, rather, that preadolescence contains signs of heterosexual interest. Pressey (52) divides childhood and youth into six periods, and mentions the emotional attachments to adults in early childhood, the tendency to one-sex groups and the clear preference for other children rather than adults in middle childhood, the one-sex groups and one-sex attachments to a chum in later childhood, and the beginning of heterosexual interests, though not centered on a single individual, in early adolescence. Cole (13) describes the early childhood fixation on an adult love-object as ending at about eight or nine; the unisexual period as intense by the age of 11 or 12; and the adult type of love-object as evident in early adolescence in the proper environment.

Elliott and Bone (17), along with several other authors, describe three stages in the growth of heterosexual love: self-love, love of one's own kind, and heterosexual love. Richmond (57) describes these stages in psychoanalytic terms: first the autoerotic stage, then the narcissistic, then the homosexual, and finally the heterosexual. She believes that the boy and girl must learn the love attitudes of their sex; that they do not achieve them without practice. She points out that few individuals achieve maturity on schedule, and that most bring with them into one stage vestiges of the preceding ones, society, as well as nature, being responsible for these delays. She believes social-sex development to be so important that it may almost be considered the equivalent of personality, "because the love attitudes extend to and color all one's human relationships." Elsewhere Richmond (55) speaks of the homosexual stage in the development of both boys and girls, describing it as the healthy beginning of acquiring kinship with one's own sex and of emancipation from the home. These stages she (55) attributes to familial and cultural, as well as biologic, influences.

Averill (1) notes the period of later childhood and pubescence as one of indifference and even aversion to the opposite sex, especially in boys. Blanchard (2) divides the normal sex life of the adolescent girl into an initial period, during which she desires to attract general attention to herself, and a later one in which she centers her attention on a particular male.

Bolton (3) writes of the period of early adolescence in boys as one in which there is an awakening of sex consciousness, but not reciprocal sex attraction; of the mutual lack of sociability between boys and girls at this period; and of the real wish to associate with the opposite sex in the middle of adolescence. Boorman (5) has an interesting diagram illustrating growth in heterosexuality, by ages (not based on actual data). Brooks (6) believes the heterosexual aspects of sex emotions are coincident with physiological maturation. Bonner (4) also describes the stages when the boy and girl have little use for each other; a second stage when each is shy and self-conscious with the other; and a third stage when there is frank interest.

Goodenough (24) describes the usual phases of social-sex development, and adds that in early adolescence the relationships of boys and girls

> retain much of their surface antagonism. They take the form of teasing, of pulling and hauling each other about, scuffling or razzing each other, and this back-handed way of showing mutual interest is particularly likely to be resorted to if observers are present.

She believes, as do several authorities, that the unisexual period is determined not by any definite sex factor, but by a difference in play activities.

Hollingworth (28) in her summary of the literature on adolescence expresses disbelief in the saltatory character of the amatory drive, but does recognize the heterosexual phase of later adolescence as different from the behavior preceding it, when "its longings for human contacts have been vague and unlocalized, its affections attachable somewhat equally to persons of either sex." In her book on adolescence Hollingworth (29) reiterates her views of the gradual expanding of the love life, finally to center on the opposite sex, and expresses a view that the four or five years following puberty are the best in which to establish heterosexuality.

McCarthy's (36) view seems to be that during middle childhood sex interests are latent and children nearly sexless in their social relationships, and that if, toward the end of this period, the sexes are drawn together at all, it is only because of social convention or interests, since primitive children do not show this tendency. (For a check on this observation, see the later summary of the anthropological literature.) McCarthy recognizes the unisexual stage just before puberty, when the distrustful attitude of boys and girls toward each other (probably due to their unequal development) is a "wise provision of nature," preventing heterosexual interest before full maturity is achieved.

Read (54) gives the usual three stages according to the Freudian classification—the infantile (the first four or five years of life), the latency period (up to puberty), and the adult period. Strain (62) describes infancy as the autoerotic and nutritive stage, with the love object self and mother; the preschool age as a neutral period sexually, with the love object playmates of either sex; the elementary-school age as a homosexual period, the love object being schoolmates of the same sex; and adolescence as a heterosexual stage. She, like many observers, has noticed the hostile period between the sexes just before maturity, and believes it to be a preventive to premature sex attractions. During preadolescence, she notes, a boy may have a friendship with a girl, but owing to the child culture of that period, which prohibits such weaknesses, he must ignore this friendship when he is likely to be observed.

Morgan (47) mentions the unisexual phase as a healthy way of weaning the child from the mother, and necessary for the heterosexual phase.

The White House Conference (70) outlines the social-sex phases of childhood

> roughly as follows: a very early period in which there is no specific manifestation of affection; next, a period in which affection goes to the parents; next a period in which it is directed to persons of the same sex; next, a period in which affection is non-specific but directed toward the opposite sex; and finally a period in which affection is specifically centered on a person of the opposite sex.

Isaacs (31) recognizes the sexual nature of the preschool children in her school by discussing at some length, and by observational material, various situations between children which show rivalry for the love of the adult, and sometimes between two children of one sex for love of one of the opposite. Cox-Miles (44), in a helpful summary of the literature on sex, points out the necessary and useful relation between the "personalization of affection" and the "specific sexual need or drive," and that the unisexual period represents a beginning of the former. She points out also the inequalities that may exist between these two in development, and the possible effect on the individual.

Zachry (72) describes as normal stages of growth the narcissistic, the unisexual (the gang age) lasting through preadolescence and early adolescence; the homosexual or "close friendship with one or two of the same sex" stage, accompanied by hero-worship; and finally the heterosexual stage.

C. *Possible Causes and Meanings of the Various Stages in Social-sex Growth*

Many philosophical explanations have been suggested for the course of heterosexual growth in children. So far as can be learned, none represents an experimental approach to the problem, except perhaps by inference. Several authors believe the unisexual phase is the result of a divergence of play interests between boys and girls at this age, and see in it no sexual implications whatsoever. There seems to be some justification for this view, for Lehman and Witty (33) found the greatest heterogeneity of interest between boys and girls between the ages of 8½ and 10½ years, and a greater similarity in adolescence than at any previous time. That certain studies (Green, Parten, Koch) have shown some evidence of unisexual interest even in the preschool period makes it appear that at least part of the basis for any unisexual group is homogeneous play interests. The gifted children study (65) also showed a heterogeneity of interests between the two sexes, but all the way from 6 to 17 years. The correlations between the interests of the two sexes ranged from .10 to .35, while the correlation between the interests of children of the same sex, whether gifted or not, was very high (.82, .83). Hildreth (27) too found a greater similarity than dissimilarity in the sports interests of the two sexes in grades 9 and 12.

Others suggest that, normally, heterosexual love must develop through the stages of a personalized affection for parent, then for peers of the same sex, and finally for peers of the opposite sex. They regard the unisexual phase as necessary, but do not actually explain why it must occur in just this fashion, except occasionally, when it is suggested that it is a wise course that nature has taken to prevent sex relations before maturity (McCarthy, Strain).

The fact that authors do not agree as to the exact time of the unisexual period makes the various explanations inexact. Bühler and others (7) would argue that the period of the deepest friendship for the same sex is in early adolescence, as shown by the frequency of the "crush." It may be that this disagreement is only a misunderstanding of terms, for almost all authors recognize the development of deepening friendship in adolescence and interest in individual personality. At the same time, it is possible that the "crush," which is probably not a usual aspect of adolescence, is merely an example of an abnormal affectional need in an individual, and not typical of an entire age group. Even the authors who have noted the homosexual phase of early adolescence could hardly deny the strong tendency to groups of like sex in the preadolescent stages.

Morgan (47) says the stages occur because somehow or other the boy finds that love for another boy is more satisfactory than love for his mother; that love for a girl is somehow more satisfactory than love for his chum; etc. Such an explanation seems hardly adequate.

Certain authors attribute the drawing apart of the two sexes just before puberty to the fact that the boy and girl achieve maturity at different stages. (Pechstein quotes Hall and his disciples as believing this.) The logic of this opinion is probably that the girl, normally thrown with her chronological peers, feels different from her boy associates when she finds herself achieving the form and functions associated with maturity, which the boys do not appear to be achieving. She therefore withdraws to her girl friends who are of like maturity. For the same reason, she would not turn to a girl friend who was noticeably lagging in attaining maturity. The boys withdraw from their more mature girl associates for somewhat the same reason.

Other authors believe the unisexual pattern to be one of play interest, and that underneath there are indications of a real interest in the opposite sex. Goodenough (24) has stated this view well. Others deem the withdrawal to members of the same sex as a necessary form of education to aid the child in identifying himself with his own sex.

While most authors assume a definite tie-up between physiological maturity and heterosexual interest, it is a subject that has not been given much experimental consideration.

Furfey (20), whose scale for developmental age includes certain aspects of social-sex development, and Rauth (52), found small but significant correlations between total score and certain anthropological measurements in adolescent boys; in preadolescent boys the correlations were not significant. The two authors reason that there is a possibility that the sex hormone has caused the common relationship between the two factors of size and social maturity. The study by Stone and Barker (61), read in abstract, shows small but significant differences between pre- and post-menarcheal groups of girls in interests of a personal, social, and vocational nature. The differences are of "the same kind and order of magnitude as those which differentiate large groups of adolescent girls differing by approximately eighteen to twenty-four months in chronological age."

The average age of puberty in boys and girls may be thought of as indirect evidence of the relationship between social behavior and sexual maturity. A graph from Shuttleworth (60), shows the variation in age of first menstruation in different geographic areas. Colder, northern climates and warmer, southern ones are associated with later menarche. The ages vary from approximately 13½ for the United States to approximately 15½ for Northern Europe and

Scotland and China and Japan. The age for the Philippines is approximately 15 years. Furfey (22, p. 144) gives a table of norms by different authors at which 50 per cent of boys had become postpubescent. For American boys these vary between 14, 14.39, and 15.25 years for the different authors, the familiar Crampton's norm being 14.39 years. Cox-Miles (44) gives an age between 15 and 16 for boys in temperate climates; for girls, between 14 and 15. Schwab and Veeder give a mean of 13-9 for boys. In China a study (67) of 1000 boys and 336 girls gave an average age of 15.48 for boys at first emission, and 15.25 for first pubic hair; the average age at first menstruation in girls was 13.56. Many authors, including Flory (19) point out the length of time required to achieve adulthood. From measurements of skeletal age Flory finds that boys do not, on the average, achieve adulthood until the age of 20; girls, until 18. From the beginning of high school on, girls are more advanced.

Brooks (6) quotes Atkinson's data on the age of sexual maturing in girls, giving a median of 13-9.6. The bases for considering the degree of maturity in boys and girls are not comparable. Comparing the Crampton and Atkinson figures, the difference in the average is somewhat over one-half year in favor of the girl. In the Chinese data, the mean difference in age between the girl's first menstruation and the boy's first emission was nearly two years. All authorities have shown the variability in onset of puberty to be large. In the Atkinson study (from Brooks), 42 per cent of the girls matured between 14½ and 17½ years. Brooks states that his data indicate nearly the same median and modal age for maturation of both sexes, but recognizes the possible difference in standards of measurement between the two sexes. Dunlap (15) remarks:

> It is popularly believed, and taught by many texts, that girls as a rule "mature" earlier than boys, but the significance of such statements is as yet conjectural. Even if it is true that as a rule, mature ova in the female appear at an earlier age than that at which mature spermatozoa appear in the male (and we do not yet know this to be true) the fact means little. "Maturity," or the more limited "sexual maturity," is a complex matter of which we know little. Girls may "mature" in some details earlier than boys; and boys may "mature" in some details earlier than girls. Emotionally it is possible that men mature earlier than women; but even of this we cannot be certain as yet.

This mass of contradictory evidence is relevant to the present study only as a reference in later discussion of differences in age norms of social-sex development in boys and girls, and its possible relation to physiological maturity. That a relation exists between social-sex life and sexual function is also believed to be proved by evidences of behavior following castration (see Furfey, *Growing Boy.*) Parshley (49) points out the instinctual nature of sex behavior:

> Sexual behavior is so patently and ineradicably biological in essence that no cultural frills can conceal the fact with approximation to completeness and no cultural restraints can do more, at best, than accomplish temporary suppression, or degree of refinement, and a variety of transmutations.

He also reports some of the literature on observations of anthropoid apes, which show some sex life previous to puberty.

D. IS THIS PATTERN A UNIVERSAL ONE AMONG ALL CULTURES?

Where significant relations have been shown between physiological maturation and social function, one might infer that in its broad outlines, at least, the social-sex pattern would be the same under different cultures, both primitive and civilized. The writings of Margaret Mead (39, 41, 42) have been searched for such patterns in the primitive groups which she has described, with much relevance to child development. In addition, the two volumes of material which she collected for the Hanover Conference on Human Relations (40) and her chapter in the *Handbook of Child Psychology* (43) have been studied.

It appears that among the Manus there is some separation of the sexes into play groups during preadolescence, but no deep-rooted antagonism between them, and they are found together frequently. When they played together it was because they enjoyed that particular game; when they separated, it was because the female activities peculiar to that culture drew them somewhat but gradually apart. The similarity to preadolescent American children who organize on sex lines and show group hostility is here illustrated: "Between the boys' groups and the girls' groups there would be occasional flare-ups, battles with sea-animal squirt guns, or swift flight and pursuit." That the preadolescent age is not entirely sexless, as some have believed, is shown in the fact that these children "very occasionally . . . united in semi-amorous play, choosing mates, building houses," etc. After puberty and before marriage the tendency is toward isolating the female until marriage, even from close friendships with contemporaries of her own sex, and so strict is this surveillance that the heterosexual tendencies of the girl are not given a chance to bloom. The boys too are thrown into single sex groups during that period, indulge in deeper friendships with their own sex, and "make more of the casual homosexuality current in childhood." In this culture adolescence is distinctly not a period of heterosexual adjustment. Ideas about sex are prudish and the adults try to stifle the sex interest of adolescents.

In Samoa, on the other hand, the period between 8 and 12 years is marked by distinct antagonism between sex groups, owing probably to incest tabus and the fact that preadolescent girls are forced to be the nursemaids of younger siblings. Girls who have reached puberty, and until they are married, do not again band together, but trust perhaps only one confidante, who aids them in their amatory adventures. In this culture adolescence is definitely a period of heterosexual adjustment, with free experimentation between boys and girls until marriage. Boys continue to band together until the age of 17. The group which is just past puberty is growing away from the sex antagonisms of the younger children, but is not yet observably sex-conscious. These children meet on formal or informal occasions and there is a great deal of good-natured banter, and even tussling, between them. The teasing is usually about some love interest in a grotesque member of the opposite sex. The picture seems very like that of American boys and girls of this age at an informal party. Earlier, between 8 and 12 years, the animosity between the sexes takes the form of attacking and fleeing from packs of children of the same age and opposite sex. Deep friendships are not common at this age, but are more so in adolescence. Among the girls even a strong friendship between an older and a younger adolescent girl never approaches the emotional intensity of an American schoolgirl "crush."

The Arapesh present a third and different picture. One gathers that they do not spend any considerable time in sex-age play groups in preadolescence. When with their contemporaries, the girls work. The boys play together to some extent, "hunting caterpillars" and the like. Friendships between children of the same sex are encouraged, and pairs of children spend much time together. Boys and girls, men and women, are not regarded as essentially different in temperament. There is an easy warmth of affection between the sexes, and, indeed, between all persons of either sex, which is not supposed to lead to sexual excitement. The heterosexual adjustment is on a different basis from that of either of the other two cultures, since the boy is affianced to a younger girl several years before marriage, and his duty is to "grow" her by feeding her and working for her welfare. The achievement of heterosexuality in this undichotomized culture seems to be a non-existent problem, for sex differences and sex excitement on a highly emotional plane do not diffuse their whole concept of personality, as they do in Western countries.

In another of the primitive societies, the Mundugumor, described in *Sex and Temperament,* preadolescent boys and girls never play with each other. The boys form in kinship groups; the girls never form play groups, but do sit quietly talking to one another. There is much strict chaperonage of adolescent girls. The children are affianced early by their families. In this culture the two sexes are antagonistic, fierce, and aggressive in marriage.

Among the Tchambuli, however, the pattern reverses our Western one. The men are the dependent sex, artistic, luxurious, and unbusinesslike, and the women are the aggressive sex, practical and in control of the purse-strings. The descriptions of the early group life of children in this culture are not as detailed, but it appears that they are identified with their own sex at seven or eight, when the boys begin to hang about the men. After the scarifying ceremony takes place, between 8 and 12, the boys are no longer identified with the women's houses, but go over to the men's.

In contrast with the Manus, who postpone heterosexual adjustments until marriage, and long after puberty, Mead mentions two societies, the Dobuan and the Trobriand, in which heterosexual play occurs in middle childhood. The Ba Thonga have a society in which preadolescent children are little outlaws and are little attended to by adults. During the season when they are supposed to be scaring the birds from the crops, the boys and girls gang together along sex lines and there is vigorous opposition between the sexes. At other times the sexes are separated in their activities. Near puberty—and here the difference from our culture may be noted—there is a change, and instead of hostility between boys and girls, they pair off, each pair building a little house together. There is an increase of sex activity between the two as time goes on. From about 10 to 15 years of age the boys are isolated from girls and sent off to a "circumcision school." Thus in the early period of adolescence there are no opportunities for heterosexual growth; instead, these have occurred earlier.

Among the Cheyenne Indians, children had friends of their own sex. Boys stayed in the children's play group until about 12, the girls until puberty, and thereafter the girls were strictly chaperoned. Mandelbaum (38) gives a description of the Bataiga, among whom girls formed friendships near puberty, preserved a good deal of propriety in the presence of boys, and were strictly chaperoned in puberty.

These pictures of primitive life, while they do not give complete descriptions of the social-sex relationships of boys and girls, do give enough information to lead one to doubt that the prevailing pattern in this country is wholly a biological one, and to suspect that society as well as the sex hormone has a part in determining the social-sex pattern of children. While there are more examples among Mead's observations of a separation of the sexes in preadolescence and of hostility, than of sympathy, at this time, a single example of a culture where the opposite of these conditions obtains is enough to vitiate the argument that the unisexual period is exclusively a matter of biological activation. In the light of some of the data from primitive societies, it is possible also to question the sexual latency of the preadolescent period. It is also clear that adolescence is not the outstanding period of heterosexual adjustment in many societies, since there are many examples of isolation during this period. Whether or not the achievement of heterosexuality is blocked by such isolation, as many authors claim, would be an interesting question for further study.

In American writings much has been said about the need for coeducation, or mixed schools, to bring about heterosexual growth. The same plea has been made by group work leaders. It is logical to assume that heterosexual growth cannot take place, as Dimock says, in a vacuum, and that opportunity must be provided for young people to meet under favorable circumstances, but there are few if any figures to prove this assumption, unless the differential marriage rate among women from coeducational and women's colleges be taken as proof. Probably more important than providing mere propinquity is the training of boys and girls throughout their educational experiences, in homes and in schools, in a normal attitude toward heterosexuality.

One group that advocates not only sex information but also the proper *milieu* and training in proper attitudes toward the other sex as the only adequate kind of sex education is the Committee responsible for the White House Conference volume, *Social Hygiene in the Schools* (70). At the preschool period, for example, they advocate training in an appreciative understanding of the family and its functions and responsibilities; at the school age (6 to 9), wholesome guided play between boys and girls and an attitude of sportsmanship; in the prepubertal period (10 to 11),

> "a respectful, courteous attitude toward children of the opposite sex, free from undue sex-consciousness, and a definite appreciation of the values of wholesome comradeship and friends and practice in securing them"; in adolescence, "freedom from restraining inhibitions, and appreciation on the part of the adult for the need for the child to achieve his heterosexuality during this time." "It is of utmost importance to use constructively the natural sex-social interests and impulses of adolescent boys and girls—their desire for one another's companionship. Such companionships should be encouraged and guided, not repressed."

They stress the need for variety in friendship as a safety factor in the lowered emotional threshold of adolescence.

Among those advocating social contact and experience with the opposite sex in recreation and school are S. L. Pressey (52), Williams (71), Cole (13), Dunlap (15), Mira (45), Hollingworth (29), and Busch (10). Hollingworth remarks that coeducation cannot produce sex attraction, nor segregation prevent it, but that the effective thing to do is to provide wholesome contacts in everyday school and recreational associations.

It is generally supposed that one-sex groups in schools provoke homosexuality, "crushes," and the like, and there is considerable evidence to support this view in our biographical and fictional arts. Cox-Miles (44) explains the probable origin of various affectional tendencies. Thus, normally the affections grow from parent to parent substitute, to a member of the same sex, to a member of the opposite sex. At any point in the series, growth may be arrested. Life circumstances are probably causative in these matters. Consider, for example, the fixation of the little girl in *La Maternelle* on her teacher at the Montmartre school. Furfey (22) points out the possibilities for arrested social-sex development at each stage of childhood.

An interesting study along these lines was made by Hurlock and Klein (30), who sent a questionnaire to groups of children, asking them about their "crushes," and to adult teachers and camp counselors who might have been the object of such "crushes." In this study the word was not limited to strong emotional attachments to persons of the same sex, as it is used in the present study. The study showed with surprising clearness that by far the greater number of such attachments had contemporaries, not teachers or counselors, as their objects. While the study showed such attachments to be common during adolescence, most had members of the opposite sex as their object. Girls were the subjects of almost all the true (homosexual) "crushes." Schools, parties, and dances were the places where such attachments most frequently developed. The authors observe that in the sense of an attachment for someone of the same sex, "crushes" are relatively infrequent, are usually for a contemporary, and are fostered when children are isolated with members of the same sex for considerable periods, as in the one-sex camp. The study does not answer the question of whether the preadolescent or adolescent period is relatively more homosexual in its tendencies, or whether the attachments they define as "crushes" are more frequent in adolescence than before. As reported, the study does not adequately show the age pattern of social-sex development, though it has other interesting implications.

That limitation of friendships to one's own sex does not necessarily imply a preference is shown in an interesting study by Stutsman (63) of two Detroit groups, one a welfare group, the other somewhat above that status. To the question, "*Do you limit friendships to your own sex?*" 44.3 per cent of the welfare group and 27.5 per cent of the other group answered "*Yes.*" However, to the question, "*Do you regard marriage as essential to future happiness?*" 44.3 per cent of the welfare group and 53 per cent of the other group answered "*Yes.*" The study showed that a small percentage of each group suffered from fear of sex inferiority, shyness, or conflict between sex and morality.

Bühler (7, 8), in speaking of the "negative phase" just preceding menstruation in girls, says it is often followed by a "crush" period, which is compensatory to the isolation of the previous period. Her study of diaries led her to believe that a flirtation or a "crush" almost always enters in (from 13-9 to 17 years) with girls, while boys more often show admiration for a leader or for a girl or an older woman. With Hurlock and Klein, she finds that the basis for these friendships is usually some particular aspect of personality which is strongly admired and emulated. Bühler quotes Danziger as finding friendships among preadolescent girls in a camp as quite shallow in comparison to those of adolescence. The concept of a frequent hero-worship in adolescence among the European writers is probably related to the "crush" of American authors.

E. WHAT DETAILED ASPECTS OF SOCIAL-SEX GROWTH, OTHER THAN REPORTS OF GENERAL PREFERENCE FOR ONE SEX OVER ANOTHER, ARE MENTIONED IN THE LITERATURE?

The literature thus far reviewed shows no attempt to determine just what aspects of conduct logically fit into the concept of social-sex growth, and it would appear from a reading of the anthropological literature that no one pattern would fit all societies. The pattern of adolescent social-sex adjustment in America is now very different from that of 50 years ago. *Middletown* (34) shows plainly the difference between the amorous behavior of young people of one generation and those of the next. An increase in "petting" in the present generation was indicated in *Middletown* and is assumed by most authors (Thom, 66). Reasons for the change are variously given as the need for sublimation because of postponed marriages, the insidious influence of the sex film, the motor car and the dark movie, preventing adequate chaperonage, etc.

Hollingworth mentions care for personal appearance and dancing as part of the pattern of beginning heterosexual adjustment in American youth. Macauley's (35) study shows the distinct difference in attitude toward clothes in the adolescent and the school-age child, the latter caring mainly for comfort and perhaps bright colors, and the former willing to sacrifice bodily comfort for style. He found from the 10th to the 12th year a beginning of attention to cut and style which marked the transition from one period to the next. Sex modesty in varying degrees is another aspect of social-sex development which sometimes forms a part of the concept. An unpublished study from the Merrill-Palmer School showed the beginning at about age 10, in a group of upper-class American children, of strong feelings about dressing and undressing in heterosexual groups of their own age, and even with adults. This pattern, according to Havelock Ellis (18) is a cultural one. Mead's studies show a great variance between the Samoan and Manu culture in this regard.

The concept of social-sex as used in this study and explained in the section on *Method* is much more detailed than in any of the studies so far met with.

III. METHOD

As already stated, the idea for this study of the development of social relations between the sexes had its origin in casual observations made in recreational groups at the Merrill-Palmer School. The various stages in this development became especially noticeable in one club that included boys and girls of a considerable range in ages, some pubescent, others who apparently would not become so for some time. The older girls of the club began to ask for dancing and some of the boys of the same age showed a willingness to join in. When dancing was begun, the prepubescent children hung back, especially the boys, who asked to return to the crafts room, made fun of the boys who were old enough to dance, and generally showed disapproval and lack of interest in an activity so clearly involving the opposite sex.

At this time a beginning of the study was made with an observation sheet of 20 statements intended to describe the various aspects of such development as they had been observed in the clubs. This sheet was checked at a Christmas party which the oldest club planned and directed, but since no definite follow-up was made, the results will not be reported here.

The study was begun along more serious lines in 1932-33, when a graduate student interested in the problem, Isabelle Unruh, was registered at the School. At this time there was an uncontrolled observation period of six weeks during which staff members and students reported instances of behavior which might enter into the concept of social-sex behavior, which by that time the author had roughly defined. These instances, plus certain aspects derived from the literature and from previous experience with children, were then classified into general statements. On the basis of this list of statements several adults were to make controlled observations on each child over a period of time. These observations, in turn, were to form the basis for a study of social-sex development of children, and a scale was to be formed from them.

A. COMPOSITION OF THE GROUP OBSERVED

With few exceptions, the children enrolled during the period of the study were graduates of the nursery school. They had been club members for varying lengths of time. The younger ones had enrolled at five years, some of the older ones not until they were 12, owing to the fact that the clubs were not begun until 1929. The group composition with respect to age, intelligence, and socio-economic status has been ascertained through a sample of the enrollment in the spring of 1932, and since the character of the group did not alter appreciably in the ensuing years, these figures may be taken as descriptive of the group throughout the period under consideration.

The total number enrolled at this time was 95, of whom observations of social-sex development were made on 85. There were more boys than girls: 46 boys, 39 girls. The mean age of the total group was 9.91 years for the boys; 9.57 years for the girls. The standard deviation in age for boys was 3.36 years; of the girls, 2.74 years; that is, there is less variation in age among the girls. The mean *IQ*s (Stanford-Binet) of the boys (117.3) and of the girls (116.5) are nearly the same, but there is greater variation among the boys (standard deviation, boys, 17.08 *IQ* points; girls, 13.89 *IQ* points). In socio-economic status, as measured by the Sims Scale, the boys and girls are nearly equal. There appears to be more homogeneity in the group in this respect than in intelligence (standard deviation, 4.43 for boys, 3.89 for girls). As a whole, the group is a privileged one. The mean (boys, 27.0; girls, 28.1) is "highest" according to Sims, ranking 9th in a scale of 10 points.

B. Program of the Clubs at Period of Observation

The program of the clubs varies according to age, and in addition is experimental, seldom remaining the same from one year to the next. An honest effort is made to make the program pleasing to the children and also to secure maximum social growth under it. That it succeeds in interesting the children is proved by the steady and voluntary attendance from year to year. The program involves a good deal of freedom of choice by each child of crafts, games, indoor and outdoor sports, dramatics, cooking, dancing, etc. The club members dine together, and leave at varying intervals afterward, depending on their age. The after-supper program for the oldest club offers a choice of dancing or games,—chiefly games for the preadolescents. In general, the plan is to have individual activities before supper and group activities afterward. A fuller description of the program has been published (11).

The children are guided by a number of adults, including the recreational clubs staff and a number of Merrill-Palmer students, each of whom is assigned to direct a special activity, such as a group in dramatics, cooking, painting, football, etc. The students do not make observational notes in the clubs as they do in the nursery school, for these older children are both very conscious and suspicious of note-taking in their presence. Most of the data taken on these children are recorded in retrospect.

A generalized description of a club program, following the activities of a single child, and showing the opportunities for mingling with or remaining aloof from the opposite sex, may be helpful in connection with the present study.

We will suppose that the child is a girl, belonging to a club in which the average ages are 11 for boys and 10 for girls. She enters the clubroom at about 4 o'clock. At the workbench is a man counselor, busy at something, and perhaps two boys, putting their last week's work in a clamp, ready for sawing. The girl, seeing the boys there, and none of her friends, looks further. At one table is seated a woman student, with leather and tools. The student asks the girl if she would like to take up leather work. The child hesitates, looks over the room, and finally consents. The student shows her the possibilities—the kinds of things she could make. The child makes her choice, gets to work, and concentrates for awhile. By this time several other girls have come in. They too drift to the leather table, seemingly because sex lines appear to have been drawn. They all work hard, talking about their school, teachers, etc. They pay no attention to the boys in the room.

Another man counselor is organizing a dramatics group, which a certain group of children always ask for. The nucleus is two girls. After the group has decided on a play, they discover they need a boy for one of the parts. They go to a "sissy" boy who is hanging around doing nothing and ask him if he will be the prince. He readily consents. He is the only boy in the club who will always consent to take part in dramatics, which at this age is considered a girls' activity. The girls in the play treat the boy taking the part of the prince without any particular attention, either amiable or otherwise. It is a democratic group except for one girl who is rather bossy. She does not dominate the group, however. The others know she sometimes has good ideas and they seem to select from her behavior what is most useful to the group.

In the meantime, a group of boys, not caring for the workbench, have decided to go out on the playing field. When they arrive they decide it will be nice to have a baseball team. They are led by a man counselor. They send a delegation back to the clubroom to ask the girls if they won't play; they need more children and neither sex has enough children to play a two-sided game. Some of the girls and boys readily consent, and one woman counselor goes along. Other girls consent rather grudgingly to go over, but make up their minds not to take part. Our little girl at the leather table shows no interest in joining the boys, but finally, since it is a fine day, the adult suggests that they all go over and help out with the game. The girl and her close friend, who has also been working with leather, at first say they don't care to, but on finding themselves deserted in the clubroom they go arm in arm over to the field.

There they find the process of choosing sides going on. Two outstanding boys have been chosen as team leaders. Each takes turn in choosing a team mate. The counselors are chosen first, without any question, and the better players among them first. Among the children, the bigger boys, especially those enthusiastic about athletics, are chosen first. The bigger girls are chosen before the smaller, less athletic boys. When the choosing is over, there are boys and girls on both sides, though there might have been more of a tendency to split along sex lines had the two captains been a boy and a girl. The girl we are observing stands with her friend over by the fence and consents to play only under pressure from the adult when more players are needed on one side. She regards herself as too old to indulge in such tomboy games, but condescendingly agrees to play if her chum will join her. She plays very well—much better than the more enthusiastic small boys.

When it is time for supper, she and her friend walk over arm in arm to the clubroom, run to the washroom, slam the door, being very particular about the possibility of any boy's peeking, and only under persuasion allow the counselor and some other little girls in, to begin washing. The two are quite thorough in their washing and comb their hair as well.

They choose a woman counselor of whom they are very fond to supervise them at table and proceed to dash over to the diningroom, seize the favored table, and put the place cards of this counselor and certain other girls they like at their table. As boys come up and try to take the table or absent-mindedly sit down, the girls are vociferous in demanding that they leave immediately. During dinner they chat about plots of movies they have seen, mostly thrillers. They eat rapidly and do not spend much time at table duties. During dessert they begin to whisper with the counselor. They are obviously plotting something. It is a plot against the boys. They will hurry through, dash back to the clubroom, and as the boys come in they will throw water on them. There is great excitement. The boys are aware that something is going on and send one of their number to spy on the girls. The girls catch him, scold him roundly, and chase him back. He returns to his friends and tells them the girls are plotting something, but gets the details wrong. The boys whisper and make a counterplot.

The dénouement is easy to imagine. The girls lie in wait, the boys sneak cautiously down, one girl makes overtures of attack. Then all scream and run, the boys after the girls, and for about half an hour, all over the block and among the bushes and around corners there is evidence of pursuit and capture. One girl is held roughly as captive and complains loudly that they are hurting her. One or two counselors who have joined in the game finally collect the group and bring them inside, where they talk excitedly of their exploits. After the girls have primped—the older ones among them have been most enthusiastic in this game—the entire group joins happily in a game of "*Simon Says Thumbs Up.*" There are two sides, with a boy and a girl for captains, and the teams are chosen on sex lines.

At the closing hour the parents call for the children, and our girl goes home in a car with her friend, her friend's father, and a boy of the same age, but less mature. They have a free-for-all tumble in the car on the way home, and though the girls pay more attention to each other than to the boy, they include him in their fun.

To continue the description of the method followed, the observations made by staff and student counselors during a six-week period were utilized by Unruh in a form containing 125 statements. After one week's use, it was revised, certain statements being altered and others added, to a total of 149, all intended to describe some aspect of social-sex development (see Table 1). Some selected graduate students and each of the staff members working with the children were then asked to fill out one of these blanks for each child at

TABLE 1

The Social-Sex Development of the Child

( ) 1. Seems to prefer to sit next to members of the same sex.
( ) 2. Seems to prefer to sit next to members of the opposite sex.
( ) 3. Will sit next to member of the opposite sex if asked to do so.
( ) 4. Dislikes to sit next to members of the opposite sex.
( ) 5. Will sit with member of the opposite sex even though he (she) is teased.
( ) 6. Will play with member of the opposite sex even though he (she) is teased.
( ) 7. Seems to enjoy physical contact of member of the opposite sex.
( ) 7*a*. Does not notice physical contact of member of the opposite sex
( ) 8. Seems to dislike physical contact of member of the opposite sex.
( ) 9. Will touch member of the opposite sex only in games or dancing.
( ) 10. Avoids physical contact with member of the opposite sex.
( ) 11. Seems embarrassed in a group consisting entirely of members of the opposite sex.
( ) 11*a*. Is not embarrassed in a group consisting entirely of members of the opposite sex.
( ) 12. Will not remain in a group composed entirely of members of the opposite sex.
( ) 12*a*. Will play in a group composed entirely of members of the opposite sex.
( ) 13. Shyness with opposite sex more apparent when with several than when with only one.
( ) 14. Seeks play in a group consisting entirely of members of the opposite sex.
( ) 15. Seems to prefer taking rôle of opposite sex in a play.
( ) 16. If asked, will take rôle of opposite sex in a play.
( ) 17. Refuses to take rôle of opposite sex in a play.
( ) 18. Makes teasing remarks about members of the same sex who are willing to take rôle of opposite sex in a play.
( ) 19. Makes teasing remarks about members of the opposite sex who are willing to take rôle of opposite sex in a play.
( ) 19*a*. Sees no incongruity in member of one sex taking rôle of opposite sex in a play.
( ) 20. Fights physically with members of the opposite sex.
( ) 21. Never fights physically with members of the opposite sex.
( ) 22. Never fights physically with members of the same sex.
( ) 23. Shames those who fight physically with members of the opposite sex.
( ) 24. Seems conscious of attention given him (her) by member of the opposite sex.
( ) 25. Tries to attract attention of members of the opposite sex, but not of same sex.
( ) 26. Accuses members of the same sex of trying to attract attention of members of the opposite sex.
( ) 27. Teases members of the same sex about attention they show members of the opposite sex.
( ) 28. Will participate in activity provided members of the same sex are included.
( ) 29. Teases members of the opposite sex about attention they show members of the opposite sex.

TABLE 1 (*continued*)

( ) 30. Will participate in a play provided members of the opposite sex are excluded.
( ) 31. Prefers company of one particular person of the same sex to the exclusion of others of that sex.
( ) 32. Prefers company of one particular person of the opposite sex to the exclusion of others of that sex.
( ) 33. Spends most of the time with members of the same sex of the same age.
( ) 33*a*. Spends most of time with members of the opposite sex of same age.
( ) 34. Carries on much secret conversation with members of the same sex.
( ) 35. Carries on much secret conversation with member of the opposite sex.
( ) 36. Plays pursuit games spontaneously in fun with members of the opposite sex.
( ) 37. Does not play pursuit games spontaneously in fun with members of the opposite sex.
( ) 38. Teases or torments members of the opposite sex by hitting back and forth.
( ) 39. Inclined to prefer members of own sex in games not involving physical skill.
( ) 40. Inclined to differentiate between sexes in games not involving physical skill.
( ) 41. In games, seems to prefer side including favorite of the opposite sex.
( ) 42. Dominates members of the opposite sex in a group.
( ) 43. Appears to acknowledge right of male to dominate; submits to boys, but not to girls.
( ) 44. Is more interested in his work than in playing with member of the opposite sex.
( ) 45. Puts down what he (she) is doing when asked by member of the opposite sex to join him (her) in work or play.
( ) 46. Seems more interested in what member of opposite sex is doing than in what member of the same sex is doing.
( ) 46*a*. Shows some interest in what members of the opposite sex are doing.
( ) 47. Shows no interest in what members of the opposite sex are doing.
( ) 48. Shows no interest in what members of the same sex are doing.
( ) 49. Makes things for member of the same sex, including adult.
( ) 50. Makes things for member of the opposite sex, including adult.
( ) 51. Enjoys associating with adult members of the opposite sex more than adult of same sex.
( ) 52. Enjoys associating with adult member of the same sex more than with adult of the opposite sex.
( ) 53. Prefers adult member of opposite sex to member of opposite sex of same age.
( ) 54. Attentive to adult members of opposite sex—serving them at table, holding doors open, etc., but not to members of the same sex.
( ) 55. Seems to prefer conversing with adult members of the same sex to members of the same age of the same sex.

TABLE 1 (*continued*)

( ) 56. Seems to prefer conversing with adult members of the opposite sex to members of opposite sex of same age.
( ) 56*a*. Seems to prefer children of either sex to adults of either sex.
( ) 57. Hero worships particular adult of the same sex.
( ) 58. Hero worships older children of the same sex.
( ) 59. Hero worships particular adult of the opposite sex.
( ) 60. Hero worships older children of the opposite sex.
( ) 61. Admires clothes of adults.
( ) 62. Imitates habits, gestures, etc., of adults.
( ) 63. Hero worships literary characters.
( ) 64. Invites adults of opposite sex to sit next to him (her), but will not invite member of opposite sex of same age.
( ) 65. Objects to having adults of either sex show signs of physical affection.
( ) 65*a*. Does not object to adults of either sex showing physical affection towards him.
( ) 66. Enjoys receiving physical affection from adults of either sex.
( ) 67. Will show affection physically with particular member of same sex of own age.
( ) 67*a*. Will not show affection physically with particular member of the same sex of own age.
( ) 68. Parent of opposite sex seems to be favorite.
( ) 69. Parent of same sex seems to be favorite.
( ) 70. Enjoys dancing with almost any member of the opposite sex.
( ) 71. Enjoys dancing with particular member of the opposite sex.
( ) 72. Seems to prefer dancing with members of the same sex.
( ) 73. Enthusiastic about dancing; prefers it to games, singing games etc.
( ) 73*a*. Dances, but does not appear to be enthusiastic about doing so.
( ) 73*b*. Often dances with members of the same sex.
( ) 74. Never dances with members of the same sex.
( ) 75. Will dance with adult member of same sex, but not with member of same sex of same age.
( ) 76. Seems to prefer to look on while others are dancing.
( ) 77. Seems to enjoy asking members of the opposite sex to dance; will cut in on others while dancing.
( ) 78. Refuses to ask members of the opposite sex to dance; will not cut in on others while dancing.
( ) 79. In choosing partner has definite preference.
( ) 80. Seems bored when dancing with member of the opposite sex.
( ) 81. Seems shy or embarrassed when dancing with member of the opposite sex.
( ) 82. Seems bored when dancing with member of the opposite sex who is shorter.
( ) 83. Seems to prefer dancing with adult members of the opposite sex to members of opposite sex of same age.
( ) 83*a*. Seems to prefer dancing with older and taller partners.
( ) 84. Avoids dancing with members of opposite sex in general.
( ) 85. Avoids dancing with particular member of the opposite sex.
( ) 86. In dancing with members of the opposite sex objects to close physical contact.
( ) 87. In dancing enjoys close contact with members of the opposite sex.

TABLE 1 (*continued*)

( ) 88. Very attentive to member of the opposite sex; waits on him (her).
( ) 89. Protective to women and girls.
( ) 90. Protective to men and boys.
( ) 91. Protective habits just beginning to show.
( ) 92. Protective habits not yet observed.
( ) 93. Protective habits evident only before an audience.
( ) 94. Often goes places (movies and parties) with members of the opposite sex.
( ) 95. Refers to corresponding with members of the opposite sex.
( ) 96. Rides in cars with members of the opposite sex.
( ) 97. Refuses to read books which are read by members of the opposite sex.
( ) 98. Enjoys books which are read by members of the opposite sex.
( ) 99. Enjoys reading romantic novels and movie magazines.
( ) *100. Sometimes talks about marriage.*
( ) 101. Reference to marriage seems embarrassing to him (her).
( ) 102. Invites members of the opposite sex to his (her) house for small party.
( ) 103. Goes voluntarily to home of member of the opposite sex.
( ) 104. Refuses to escort member of the opposite sex home.
( ) 105. Escorts members of the opposite sex home voluntarily.
( ) 106. Will escort member of the opposite sex home if asked to do so.
( ) 107. Discusses acquaintances of the opposite sex who are not members of the group.
( ) 108. Refers to kissing members of the opposite sex.
( ) 109. Seems interested in being introduced to members of the opposite sex.
( ) *110. Seems interested in being introduced to members of the same sex.*
( ) 111. Avoids being introduced to members of the opposite sex.
( ) 112. Avoids being introduced to member of the same sex.
( ) 113. Attempts to improve his (her) personal appearance in order to attract members of the opposite sex.
( ) 114. Careless about personal appearance even when with members of the opposite sex.
( ) 115. Remarks to members of same sex about attractiveness of members of the opposite sex.
( ) 115*a*. Makes deprecatory remarks about members of the opposite sex.
( ) 116. Remarks to members of opposite sex concerning his (her) attractiveness—sarcasm, perhaps.
( ) 116*a*. Avoids references to attractiveness of members of the opposite sex although probably conscious of it.
( ) 116*b*. Not concerned with attractiveness of members of the opposite sex—probably not conscious of it.
( ) 117. Not embarrassed at reference to articles of clothing of opposite sex.
( ) 118. Seems embarrassed at reference to articles of clothing of opposite sex.
( ) 118*a*. Likes to wear clothes suggestive of the opposite sex.
( ) 118*b*. Not embarrassed to wear clothes suggestive of the opposite sex.
( ) 118*c*. Dislikes to wear clothes suggestive of the opposite sex.
( ) 119. Does not wear clothing suggestive of the opposite sex.
( ) 120. If boy, definitely seeks to wait on girls.

TABLE 1 (*continued*)

( ) 121. If girl, expects to be waited on by boys.
( ) 122. Tells others about members of the same sex who like members of the opposite sex.
( ) 123. Expresses desire to be member of the opposite sex.
( ) 124. Definitely classifies work to be done according to sex—housework to girls, moving furniture to boys, etc.
( ) 125. Does not differentiate work according to sex.
( ) 126. Definitely classifies games according to sex.
( ) 127. Does not differentiate games according to sex.
( ) 128. Shows interest in peeking at opposite sex in bathroom, dressing, etc.
( ) 129. Is flirtatious or coy.
( ) 130. Dramatizes sex; if boy, a great big he-man tough; if girl, is feminine, dainty, etc.
( ) 131. Is modest; careful how he (she) sits, dresses, etc.
( ) 132. Not yet modest as to posture, gesture, clothing, etc.

each meeting until there were four for each child filled out by the same observer. This requirement was fairly well met, though there were some exceptions, owing to absences.

The blanks were filled out after the club meeting or on the morning after, when the observer read through each statement, checking those that described some behavior observed in the child. In view of later discussion of disagreements among the raters, it should be pointed out that various circumstances might cause observers to check statements for a child indicating different patterns of behavior on the same day. The children are scattered in their activities and the individual counselor might therefore see the individual child for only a limited time and in a single activity. These limitations were inherent in the program and could not be avoided.

The observations made during the six-week period of the first year were used in part in the Unruh (68) study. Once a year during the next two years paired observations on each child in the clubs were made by two staff members. During the first years the raters were Unruh (*I.U.*) and the author (*E.H.C.*); in addition, the ratings made by student counselors and other staff members are used in part of the study. *J.M.* and *E.H.C.* were the raters the second year, and *J.M.* and *J.D.* the third year. In addition, a few children attending the Merrill-Palmer Camp one summer were rated by counselors. The raters chosen were those who were most experienced in observing children of these ages, and who, in addition, could spend the time entailed in checking the blanks.

Unruh's study reveals a difficulty in statistical interpretation occasioned by the method of rating. In her study she sought to shorten the original list of statements by eliminating statements infrequently used or unreliable. She calculated for each statement, for boys and girls separately, the percentage of the total number for which a statement was checked. However, the meaning of a checked statement is ambiguous, for it may be checked or not for various reasons. Thus, not all statements, for example those concerned with dancing, are applicable to all ages. Further, even if the statement may be applicable to the given age, and to the individual child rated as well, it may not be checked because the observer has had no opportunity to see the type of behavior in question. Thus the method of allowing the rater to check only those statements of which he feels competent to judge and to omit all others, while very satisfactory to the rater, increases the difficulty of interpretation and statistical treatment. This method is patterned after that used in the Merrill-Palmer personality scales (58), though it differs in one important respect; that is, the statements were to be scaled, not by judges, but by the results obtained by administering the test to a number of children of different ages and then scaling each statement according to its power to differentiate among children of various ages. After calculating the percentages described, Unruh determined the probable errors of these percentages and their ratios to the percentages. When this ratio was greater than four, the statement was retained as being valid for the scale; when it was less than four the statement was discarded as not sufficiently used. In this way 19 statements were discarded, leaving 130 of the original 149.

Unruh determined the reliability of the scale by correlating the paired ratings of the two observers. Tables were set up showing the presence or absence of a check for each rater on each statement. Chance might determine one-fourth of the cases in each quadrant of the table. Therefore, when the tallies in each quadrant amounted to one-third or more of the total, indicating this extent of agreement between the raters, the statement was retained. By this analysis, 73 more items were eliminated, leaving 57.

This method is not quite fair to the data, since, as already pointed out, the fact that one rater observed a certain type of behavior while the other did not does not necessarily denote disagreement, but in many cases only that one rater had an opportunity to see a child in a certain situation while the other did not. In some cases two contiguous statements show the extremes of one kind of behavior, or a trio shows the two extremes and the mean between. Actual disagreement would be shown if each rater was required to check one of the pair or the trio of statements. In some instances where the raters have checked contrasting statements such disagreement is clear even in the present scale, but unless such checking is required this method of determining disagreement cannot be used throughout the study.

Unruh reduced the statements further by an analysis of the agreement between the two raters by means of the Kelley coefficient of contingency formula. When the relationship of C to the PE_C was three or more, the statement was retained as reliable. After this analysis, 32 statements remained, of which 12, or about 37.5 per cent, concerned behavior observable only in adolescents, while this was true of only about one-fifth of the original statements. She then divided the children into three large age groups—*Middle Childhood*, ages 5.5 to 9.4 years; *Later Childhood*, ages 9.5 to 12.4 years; and *Early Adolescence*, ages 12.5 to 15.7. She determined the percentage of boys and girls who had been checked as showing the behavior described in each of the 32 statements, and from these figures made an interesting qualitative comparison of the changes in behavior in each of the details of social-sex behavior through the three periods of childhood.

The Unruh study has provided an invaluable groundwork for the present study. Unruh suggests that her study be extended in order to develop a total scoring method for the scale, and that further data be collected. In the present study the author, believing that the statistical methods used in the Unruh analysis destroyed some opportunities of studying the more detailed aspects of social-sex development, especially by eliminating statements applicable to the lower age groups, has returned to the original list of 149 statements. A different method of eliminating statements has been devised, a method of scoring the statements by age worked out, and sequential data on the same group of children have been collected over a period of three years. The social-sex development of the child is described on the basis of these results.

C. SELECTION OF STATEMENTS

A code number, to be used in the analysis instead of a name, was assigned to each child. All data for the three years were then divided into two groups by sex. A tabulation was then made for each separate statement, showing on the horizontal the name of each rater who had taken part in the study, and on the vertical, the entire distribution of ages in the study *by six-month age groups*, beginning at 5-9 and ending at 18-5. In each square thus formed were placed the code numbers of children of that age who had been marked as showing the trait indicated, and by the rater indicated. For example, for the *CA* group 6-0 to 6-5, and for the statement 132, *"Not yet modest as to posture, gesture, clothing, etc.,"* it is shown that Boy 20 has been marked as having this trait by raters *J.M.* and *J.D.*, who made simultaneous ratings of him at this age. A year earlier raters *E.H.C.* and *J.M.* had rated Boy 20 as showing the same trait. Boy 20 does not appear in later ratings; thus the rating at age 6-0 to 6-5 is the latest we have.

At the left of these sheets, covering the data for each statement is a listing of the relative frequency of marking of each statement, as follows: First, the total number of children of the given age and sex marked by at least one rater as having the trait indicated, and the percentage of the total number of available children of this age and sex represented by this number; second, the total number of ratings of the trait for the age and sex and its percentage of the total possible number of ratings.

These two percentages are quite different and are based upon different criteria. The first disregards the total number of identical ratings given a child by two or more raters, and depends rather upon the powers of observation of a single rater. The second disregards the single rater's powers of observation and depends upon the number of raters who have marked a trait at the age indicated. Theoretically, the second method appears to be the sounder, but actually, under the conditions of observation, the first is more so, owing to the different meanings of the unchecked statement. As already explained, the fact that one of a pair of raters checks a statement for a child while the other does not, does not necessarily imply either absence of the trait or disagreement, but merely that one of the raters may have had an opportunity to observe it while the other did not. Thus, in the present study there was not the same opportunity for agreement or disagreement between the raters which obtains in some experimental situations. From three years' work with the scale, the author believes that considerable importance should be attributed to the checked item of a single rater on a single child.

The next step was making a series of data sheets from the original sheets, using various combinations of age groups for each statement by sexes. The result was data on percentages of marked statements by sexes for 6-month, 1-year, 2-year, and 4-year age groupings.

A word should be said about the basis of figuring the percentages for age intervals larger than the original 6-month groupings. Children were rated at about 1-year intervals. Thus the same child would appear twice in the same grouping of larger than one year. Thus, in the 2-, 3-, and 4-year intervals the same child might be counted twice or oftener. This is not the usual practice, and in the end only data based on 1-year age intervals were used in selecting statements. However, in experimenting with the various age distributions, it seemed desirable to count ratings on one child separately, both because at some ages we could not afford to sacrifice any cases, and because we believed that in a 2-year period enough growth would occur so that sequential measurements on a child in that time would add more to our scaling than would the average of these measurements.

Graphic representations of the data on 6-month to 4-year age groupings were made to show the age distribution of each statement by sex and according to coarse and fine age intervals, and these graphs were studied to determine which statements seemed to show a differentiation with age and which ones were used often enough to be included in the final scale. Thus popularity and selectivity were two of the criteria used to shorten the scale.

In studying the graphs it is interesting to see that the distributions fall in several distinct classes. In statements which describe the behavior of young children but almost never the behavior of adolescents, e.g., "*Not yet modest as to posture, gesture, clothing, etc.*," the percentage falls from a high level at the youngest age to a low level at adolescence, producing a descending curve. In the case of statements describing behavior rare in young children but almost universal in adolescents, e.g., "*Attempts to improve personal* *appearance in order to attract members of the opposite sex,"* we have an ascending curve. There is also a group of statements almost universally true of middle childhood, but almost never true of younger children or adolescents, and producing, in graphic form, an inverted *"U."* Statements referring to strong unisexual feelings fall in this group, e.g., *"Seems to prefer to sit next to members of the same sex."* Statements of the opposite sort, e.g., *"Is not embarrassed in a group consisting entirely of members of the opposite sex,"* produce a *"U"* when presented graphically (see Figure 1).

The peaks of similar curves did not fall at the same age for every statement, and the curves for boys and girls were also different in this respect. The curves differentiating by age and those not differentiating by age both stand out clearly.

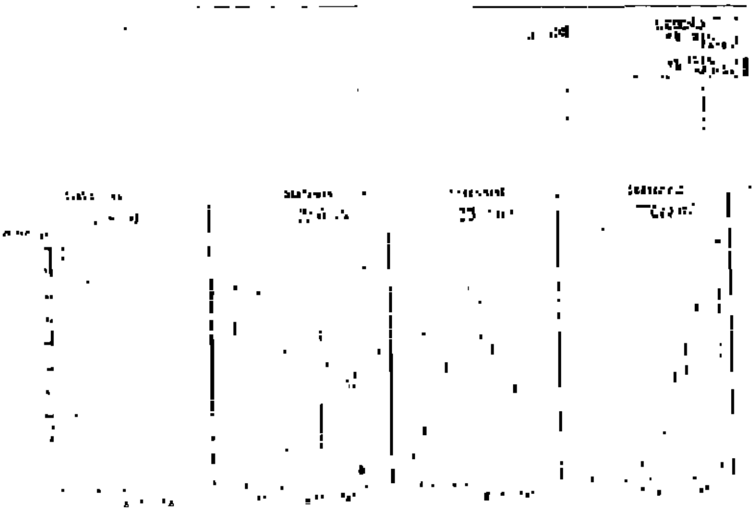


FIGURE 1

Graphs Showing (*a*) Descending Curve, for Statement 7*a*, Characteristic Form of Curve Assumed by Distributions for Statements Describing Social-Sex Behavior Appearing Most Commonly in Early Childhood; (*b*) *U*-shaped Curve Statement 12*a*, Characteristic for Early and Late Childhood; (*c*) Inverted *U*-shaped Curve, Statement 53, Characteristic for Middle Childhood; (*d*) Ascending Curve, Statement 70, Characteristic for Later Period

Several ways of selecting curves for use in the final revision were considered. The method adopted was as follows: Curves based upon a one-year age interval were used as a basis for judgment. Curves based upon a coarser age interval were studied when the pattern was not altogether clear from the former curves. The curves based upon the number of children regardless of the number of raters, rather than upon the number of children times raters, were used. *Statements were omitted when the peak of the curve* did not reach a height of 50 per cent; those which satisfied this criterion but whose lowest point was not below the height of 50 per cent were also omitted. The reasoning followed was that the behavior in question should be common enough at one age so that more than half the children had exhibited it, as observed by the rater, yet infrequent enough at other life periods so that less than half the children exhibited it. The curves based upon the percentage of the total number of children rather than the percentage of the total number of children times raters were selected as a criterion because they showed greater contrasts between ages in all statements.

A few more statements were eliminated when it was found that two similar statements had been included and that one could very well be omitted, and a few because the age location of the peak point was difficult to determine, owing to an atypical curve.

The purpose of the statistical process up to this time was to ascertain whether a scale value could be assigned to each of the statements on the basis of age, so that a total score could be secured. By means of such a score, it was thought, the agreement between raters and the social-sex growth of individuals from year to year could be *studied quantitatively.*

At about this time Dr. Furfey suggested to the author that a fairly coarse scoring system would prove to be just as effective as a fine one, and far easier to manipulate statistically. The author had intended to find a "social-sex age" value for each statement and use this for the scale value. Dr. Furfey suggested using a three-step scale instead, with a value of zero for curves that sloped downward steeply to the right (characteristic of the youngest children in our groups), a value of one for those resembling an inverted *U* (characteristic of the middle ages of our groups), and a value of two for those sloping upward steeply to the right (characteristic of our adolescent children). The only curves this plan did not provide for were those reaching a peak not distinctly in the middle of any one of these three periods, and those shaped like a *U*. Provision was made later for both of these contingencies.

Both the age method and the three-step method of scoring were used at first, to test their relative value. Each graph was scanned to determine the age at which the trait in question reached its maximum height, and the appropriate social-sex age value was assigned. The same curve was given a score of zero, one, or two, depending on its type. In the case of *U* curves—statements 11*a*, 12*a*, 40*a*, and 127, for both sexes—the phrase *"because not sex-conscious"* was added, and the lower age score only was assigned. In the final revision a similar statement with the phrase *"because sex-conscious"* added will be used, supplementing each of these statements, and to these new statements the higher age score will be assigned. The observer will thus have to decide, in observing such behavior, whether the child shows heterosexual interests because he is sex-conscious or because he is not yet so. This may appear confusing as it is described, but in the actual situation of working with school-age children it is easy to separate the two kinds of motives. In the present study, however, we gave these statements, if marked positively, the lower value if the child fell into the less mature group, and omitted the statement if he fell into the more mature group.

In curves where there was a distinct peak just between the middle and the upper or lower age range, a scale value in decimals, between the proper steps of the three-point scale, was assigned. This method made the scoring more elaborate than was planned, but the simpler plan did not appear to do justice to the data.

When the statements finally selected had been assigned scale values, they were listed according to these values from low to high, or immature to mature, in order to determine how regular the steps in the scale were and whether certain areas had many more statements than others. For each sex, 68 statements remained. These were distributed in six-month, social-sex age groupings, by three-point scale values. A distinct overcrowding of statements at the adolescent age was then evident. An attempt was made to equalize approximately the number of statements for the three stages, by eliminating at the upper end every statement which could be dispensed with by any reasonable criterion. The final number of statements remaining for each age period was as given in Table 2.

TABLE 2

	Boys	Girls	Scale value
Younger period	14	11	0 — 0.9
Middle period	9	13	1 — 1.9
Older period	20	20	2 — 2.9
Total	43	44	

Apparently social-sex behavior becomes so much more conspicuous in adolescence that its various phases require a greater number of descriptive statements, though it would seem possible in a perfect scale to give equal weight to the manifestations occurring in early and middle childhood. It should be noted in this connection that the scale begins at five years and that the averaging process has resulted in making the scale values as of the middle of this younger period, rather than the beginning, whereas if three- and four-year-old children had been included it is possible that there would be more items describing the five year old.

The next step was to test the relative merits of the age and three-step methods of scoring. All blanks were scored according to both plans. The method in each case was to assign the proper scale value to each checked statement and to calculate the mean of these checked statements for each child. These scores, the name of the rater, and the chronological age of the child in years and months were recorded on permanent individual record cards. Correlations between the paired ratings of raters *I.U.* and *E.H.C.*, without holding age constant, were, for the age scoring method, $r = .847$; for the three-step scoring method, $r = .862$. In view of the close correlations for the two methods, and the fact that the three-step method made scoring easier, while the age method was perhaps a finer measure than the data warranted, the three-step scoring method was adopted, with the modifications already described. The final scales, for boys and girls, with the final scale values for each statement, are shown in Tables 3 and 4.

TABLE 3

THE SOCIAL-SEX DEVELOPMENT OF THE *Boy*

Place a check beside every statement describing behavior which you have observed to be characteristic of the child. This scale is to be used only after direct observation of children in groups.

Scale value		*SSA*	
0 - 6	( )	12*aa*	Will play in a group composed entirely of girls because not yet sex-conscious.
0 - 7	( )	11*aa*	*Is not embarrassed in a group consisting entirely of girls because not yet sex-conscious.
0 - 7	( )	65*a*	Does not object to adults of either sex showing physical affection toward him.
0 - 7	( )	132	*Not yet modest as to posture, gesture, clothing, etc.
0 - 7.5	( )	127	Does not differentiate games according to sex.
0 - 7.5	( )	92	Protective habits toward women and girls not yet observed.
0 - 7.5	( )	40	In choosing groups for games not involving physical skills, draws no sex lines because not yet sex-conscious.
0 - 7.5	( )	20	Fights physically with girls.
.5- 8	( )	7*a*	*Does not notice physical contact with girls.
.5- 8	( )	114	Careless about personal appearance even when with girls.
.5- 8	( )	125	Does not differentiate work according to sex.
.5- 8.5	( )	116*b*	*Not concerned with attractiveness of girls—probably not conscious of it.
.5- 8.5	( )	53	Prefers adult women to girls of own age.
.5- 8.5	( )	44	*More interested in his work than in playing with girls as such.
1 - 9.5	( )	36	Plays pursuit games in fun with girls.
1 - 9.5	( )	47	*Shows no interest in what girls are doing.
1 -11	( )	39	*Inclined to prefer boys in games not involving physical skill.
1 -11.5	( )	1	Seems to prefer to sit next to boys.
1 -11.5	( )	28	Will participate in an activity provided boys are included.
1.5-12	( )	9	*Will touch girls only in games or other conventionalized situations.
1.5-12.5	( )	12	Will not remain in group composed entirely of girls.
1.7-14	( )	122	Tells others about certain boys who like certain girls.
1.7-14	( )	131	*Is modest; careful how he sits, dresses, etc.
2 -14.5	( )	116*a*	Avoids references to attractiveness of girls, although probably conscious of it.
2 -14.5	( )	113	*Attempts to improve his personal appearance in order to attract girls.
2 -15	( )	24	Seems conscious of attention given him by girls.
2 -15	( )	56*a*	Seems to prefer children of either sex to adults of either sex.
2 -15.5	( )	13	Shyness with girls more apparent when with several than when with only one.

TABLE 3 (*continued*)

Scale value	*SSA*			
2	-15.5	( )	65	Objects to having adults of either sex show signs of physical affection.
2	-15.5	( )	70	Enjoys dancing with almost any girl.
2	-15.5	( )	77	*Seems to enjoy asking girls to dance; will cut in on others while dancing.
2	-15.5	( )	87	In dancing, enjoys physical contact with girls.
2	-15.5	( )	89	Protective and attentive to women and girls; waits on them.
2	-15.5	( )	124	Definitely classifies work to be done according to sex—housework to girls, moving furniture to boys, etc.
2	-15.5	( )	73	†Enthusiastic about dancing; prefers it to games, singing games, etc.
2	-16	( )	5	Will sit with girls even though he is teased.
2	-16	( )	25*a*	Tries to attract attention of girls especially.
2	-16	( )	45	Puts down what he is doing when asked by a girl to join her in work or play.
2	-16	( )	62	Imitates habits, gestures, etc., of men.
2	-16.5	( )	2	*Seems to prefer to sit next to girls.
2	-16.5	( )	7	Seems to enjoy physical contact with girls.
2	-16.5	( )	46	Seems more interested in what girls are doing than in what boys are doing.
2	-17	( )	43	Appears to acknowledge right of the male to dominate; submits to boys but not to girls.
2	-17	( )	11*ab*	Is not embarrassed in a group composed entirely of girls because he is definitely attracted to them.
2	-17	( )	12*ab*	Will play in a group composed entirely of girls because he is definitely attracted to them.
2	-17	( )	42*a*	In choosing groups for games not involving physical skills draws no sex lines because definitely interested in being with girls.

*Statements marked with an asterisk also appear in the Unruh scale.

†Numbers following parentheses are those of the statements as originally numbered.

TABLE 4

THE SOCIAL-SEX DEVELOPMENT OF THE *Girl*

Place a check beside every statement describing behavior which you have observed to be characteristic of the child.

Scale value	*SSA*			
0	- 6	( )	12*a*	Will play in a group composed entirely of boys because not sex-conscious.
0	- 6.5	( )	40	In choosing groups for games not involving physical skills, draws no sex lines because not yet sex-conscious.

TABLE 4 (*continued*)

Scale value	*SSA*		
.5- 7	( )	132	*Not yet modest as to posture, gesture, clothing, etc.
.5- 7	( )	11*aa*	*Is not embarrassed in group consisting entirely of boys because not sex-conscious.
.5- 7.5	( )	7*a*	*Does not notice physical contact with boys.
.5- 7.5	( )	44	*Is more interested in her work than in playing with boys.
.5- 7.5	( )	66	Enjoys receiving physical affection from adults of either sex.
.5- 8	( )	114	Careless about personal appearance even when with boys.
.5- 8	( )	116*b*	*Not concerned with attractiveness of boys—probably not conscious of it.
.5- 8	( )	125	Does not differentiate work according to sex.
.5- 8	( )	127	Does not differentiate games according to sex, because not sex-conscious.
1 - 8.5	( )	47	*Shows no interest in what boys are doing.
1 -10	( )	12	Will not remain in a group composed entirely of boys.
1 -10	( )	39	*In choosing groups is inclined to prefer girls in games not involving physical skill.
1 -10	( )	53	Prefers men (counselors) to boys of her own age.
1 -10.5	( )	1	Seems to prefer to sit next to girls.
1 -10.5	( )	28	Will participate in an activity provided girls are included.
1 -10.5	( )	64	Invites men to sit next to her but will not invite boys of her own age.
1 -11	( )	9	*Will touch boys only in games or other conventionalized situations.
1 -11.5	( )	126	Definitely classifies games according to sex.
1.5-12.5	( )	34	Carries on much secret conversation with girls.
1.5-13	( )	13	Shyness with boys more apparent when with several than when with only one.
1.5-13	( )	73*b*	Often dances with girls.
1.5-13.5	( )	116*a*	Avoids references to attractiveness of boys, although probably conscious of it.
2 -14	( )	131	*Is modest, careful how she sits, dresses, etc.
2 -14.5	( )	24	Seems conscious of attention given her by boys.
2 -14.5	( )	61	Admires clothes of women.
2 -14.5	( )	73	*Enthusiastic about dancing—prefers it to games, singing games, etc.
2 -14.5	( )	113	*Attempts to improve her personal appearance in order to attract boys.
2 -15	( )	5	Will sit with boys even though she is teased.
2 -15	( )	43	Appears to acknowledge right of the male to dominate; more likely to submit to boys than to girls.
2 -15	( )	45	Puts down what she is doing when asked by a boy to join him in work or play.
2 -15	( )	70	Enjoys dancing with almost any boy.
2 -15	( )	124	Definitely classifies work to be done according to sex—housework to girls; moving furniture to boys, etc.

TABLE 4 (*continued*)

Scale value	*SSA*			
2	-15.5	( )	2	*Seems to prefer to sit next to boys.
2	-15.5	( )	25	Tries to attract attention of boys especially.
2	-15.5	( )	46	Seems more interested in what boys are doing than in what girls are doing.
2	-15.5	( )	62	Imitates habits, gestures, etc., of women.
2	-15.5	( )	67	Will show physical affection for a particular girl of the same age.
2	-15.5	( )	74	No longer dances with girls.
2	-15.5	( )	83*a*	*Seems to prefer dancing with older and taller male partners.
2	-16	( )	7	Seems to enjoy physical contact with boys.
2	-16	( )	94	Often goes places (movies and parties) with boys.
2	-16	( )	115	Makes remarks to girls about the attractiveness of boys.
2	-16	( )	11*ab*	Is not embarrassed in a group consisting entirely of boys because she is definitely attracted to them.
2	-16	( )	12*ab*	Will play in a group composed entirely of boys, because definitely attracted to them.
2	-16	( )	40*a*	In choosing groups for games involving no physical skill does not draw sex lines because she is definitely interested in being with boys.
2	-16	( )	127*a*	Does not differentiate games according to sex, because she is definitely interested in being with boys.

---

*Statements marked with an asterisk also appear in the Unruh scale.

†Numbers following parentheses are those of the statements as originally numbered.

IV. RELIABILITY OF THE SCALE

A statistical analysis has been made of the scores of individual children, showing how the scores of the *various raters* compare and how the averaged scores change with age. The correlation of .30 to .40 between the paired raters, with age held constant, indicates a degree of agreement which compares favorably with that reported for paired ratings in other studies. However, the acid test of such agreement is not in the scatter diagram but in the study of the sequential records of individual children according to different raters. In graphing scores by different raters over a period of time it is difficult to tell how far real growth, and how far merely a fluctuating and unreliable observation in the *rater*, is *represented*. In the present scale there is an advantage in the fact that the scores are based upon age, and should therefore increase with age, thus furnishing some basis for judging the adequacy of raters' observations.

A brief but interesting comparative analysis was made of the raw scores given by the paired raters. Graphs were first made for the scores of 52 children who had been rated in social-sex development for three successive years. Samples of these graphs are shown

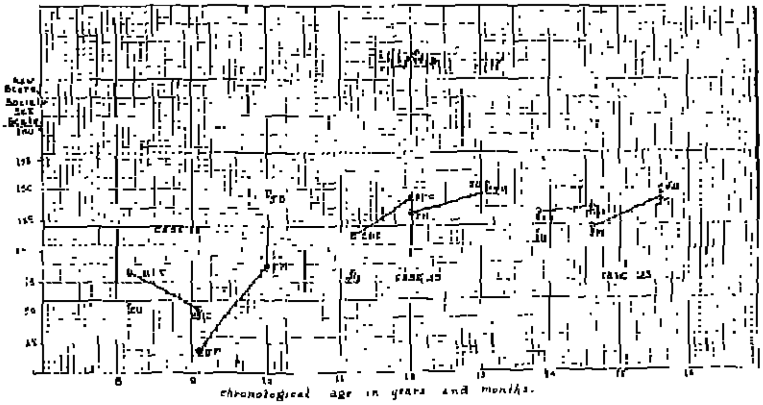


FIGURE 2

Graph Showing Raw Scores in Social-Sex Development Given by Different Raters, with One Case Showing Disagreement between Raters and Loss of Score, and Two Cases Showing Agreement between Raters and Normal Increase of Score with Age

in Figure 2. In each of the three years o
from the preceding year was carried over.
first year were *I.U.* and *E.H.C.;* for the s
J.M.; for the third year, *J.M.* and *J.D.* T
the raters were as given in Table 5.

TABLE 5

	raw r
E.H.C. and I.U.	.852
E.H.C. and J.M.	.862
J.M. and J.D.	.809

It is interesting to compare the correlat
Merrill-Palmer preschool personality chart,
quired to mark a child plus or minus on 1
In that case the raw r of scores given by a
the r with age held constant, .697. While t
this score little and that for social-sex mucl
ably to be explained by the fact that the
for age, while the other is not. The parti
raters on the social-sex scale are lower thaı
on the personality chart, showing the grea
and perhaps the greater reliability of the pe

The scores given by the paired raters
measuring the distance, on graph paper, l
senting their scores and comparing the ave
for the three pairs. These figures represent
agreement between raters. The first two
other in about the same amount; the third
degree. The average difference in scale p
and 25.02, respectively, for the three pairs

These same distances were then compare
there was a tendency toward greater disagı
age group of children. There was least di
dren aged 13 and above, next least for tl
most for the youngest, aged 5 to 8. The
16.93, 23.15, and 25.65 scale points, respe

A test was then made to determine th
ratings made by *E.H.C.* and *J.M.* which sh

with time. The reasoning was that an individual child would not be expected to lose maturity in social-sex development with increasing age, though he might remain stationary, and that downward trends would thus reveal faults either in the method itself or in the rater. An arbitrary classification of curves was made. Those changing in either direction 10 score points or less in the year were regarded as stationary; those showing ascending directions greater than 10 points as on the increase, and therefore logical; those showing descending directions greater than 10 points as showing a loss, and therefore illogical.

Rater *E.H.C.* had marked 40.38 per cent of the 52 individual curves in an ascending direction; 38.46 per cent as remaining in relatively the same position; and 21.15 per cent as descending. Rater *J.M.* had marked 46 per cent of 50 curves as ascending, 46 as remaining relatively the same, and 8 per cent as descending. These figures suggest that rater *E.H.C.* made relatively more errors of judgment of this kind. Counting changes of both increase and decrease of whatever size, *E.H.C.* shows 67.31 per cent of increase, or logical change, and 30.77 per cent of decrease. The corresponding figures for *J.M.* are 72 per cent and 20 per cent, which make the better showing. The self-correlations of raters *E.H.C.* and *J.M.* also indicate the greater consistency of *J.M.* The self-correlations were as given in Table 6.

TABLE 6

	raw *r*	*N*	*r* age held constant
E.H.C. (1933-34)	.851	59	.473
J.M. (1934-35)	.944	41	.683

The averaged scores of paired raters were also analyzed to determine whether, of two raters marking in fairly close agreement, there was a consistent tendency for one of the pair to mark higher than the other. In the pair *I.U.* and *E.H.C.*, *E.H.C.* marked 50 per cent of the cases higher than *I.U.*, 46 per cent lower, and 4 per cent the same; that is, there was no consistent trend. In the pair *E.H.C.* and *J.M.*, *E.H.C.* marked 60 per cent higher, 35.6 per cent lower, and 4.4 per cent the same as J.M. Here there seems to be a more nearly consistent bias toward higher ratings by *E.H.C.* In the pair *J.M.* and *J.D.*, *J.M.* marked 36.8 per cent higher, 57.9 per cent lower, and 5.3 per cent the same. Here, as in the first pair, there is no consistent bias, though there is a tendency toward lower ratings. There is some evidence that *J.M.* shows a consistent bias toward lower ratings than her partner in the two different pairs. Rater *E.H.C.* tends to mark higher than *J.M.*, but in no consistent direction away from *I.U.*

The matter of constant bias is interesting to study in judging the adequacy of raters and in determining the "true" score. If one rater tends to be consistently higher than her partner, one can with some confidence take a mean figure, half way between the two, and be fairly certain that the growth change measured in the resulting figure is a true one. However, if there is no constant relationship, but one rater tends to be variably higher and lower than the other, the mean cannot be taken as showing such a true change. The same difficulty obtains when at least one of the two raters changes each year. In this study, the fact that rater *J.M.* tends to rate lower than *E.H.C.* or *J.D.* in about two-thirds of the cases seems to show a true bias. However, whether it would be warranted to recognize this bias statistically and present the average between *J.M.* and her partner as approximating the child's true status is questionable, since in at least one-third of the cases the assumption does not hold.

Since the present method depends so largely upon the accuracy of adults' observations of children, the findings on the reliability of raters are very important. The following summary shows the degree of adequacy of the raters in the present study and may possibly suggest criteria for other studies in which a combined rating-observational technique is employed.

1. It appears that agreement between raters is somewhat dependent upon educating raters to the meaning of the concepts implicit in the scale. A rater experienced in working with and observing children, but without such a specific understanding, produces ratings less in agreement with those of an experienced partner than are those of a rater who, with equal general experience, has given more thought to the concept of social-sex development.

2. It is easier to reach an agreement in the rating of adolescents than in rating children of the prepubescent stage (about 9 to 12), and easier, in turn, to agree upon the latter than on younger chil-

dren (5 to 8). This result would be the expected one, since much of a younger child's social-sex behavior is shown negatively rather than positively. However, in learning to observe the various stages of the development of heterosexuality, it is probably just as important to be able to recognize signs of normality in one period as in another. Thus, in instructing raters, special emphasis should be laid upon the phases of social-sex development manifest in younger children.

3. The following criterion, based upon the results of the best rater in the present study, may be used in judging the efficiency of a rater. Not more than 10 per cent of the second year's ratings of a child should fall more than 10 scale points below those of the first year; not more than 20 per cent should be lower to *any* degree. Between 40 and 45 per cent should increase more than 10 score points, and at least 70 per cent should increase by some amount, large or small.

4. In the present study, the ratings of one reliable rater proved more satisfactory than the average of paired ratings. This conclusion was reached after a study of the reliability and direction of bias of the various raters, and the constancy of relationship between pairs of raters.

MEAN SCORES AND PERCENTILES FOR THE SCALE

Mean scores for boys and girls for ages from 5 to 16 or 18, by two- and four-year intervals, are shown in Table 7.

TABLE 7

SOCIAL-SEX DEVELOPMENT: MEAN SCORES FOR BOYS AND GIRLS BY AGE

Age	Boys Number	Boys Mean score	*SD*	Girls Number	Girls Mean score	*SD*
			Two-year interval			
5- 6	24	53	19.00	16	67	11.73
7- 8	21	64	20.41	24	78	16.06
9-10	20	83	24.37	26	94	13.84
11-12	20	107	26.80	16	110	24.08
13-14	13	140	21.42	18	154	24.51
15-16	10	172	14.21	7	182	9.56
17-18	7	169	11.62	—	—	—
			Four-year interval			
5- 8	45	56.4		39	73.3	
9-12	38	93.1		40	100.9	
13-16	23	153.8		25	161.8	

The scores of the girls are more advanced than those of the boys at all ages. The variability of the various age groups is illustrated in Figures 3 to 6. The boys are more variable than the girls in the first two age groups. In both sexes there is a considerable overlap in the two- and four-year intervals. That the overlap in the 13 to 16 age group is chiefly due to the great variability of children aged 13 and 14 and very little to those aged 15 and 16 is indicated in Figures 5 and 6, showing distributions based on two-year intervals. In both sexes there is a wide spread for ages 11 and 12, as well as for ages 13 and 14, with little spread for ages 15 and 16. Perhaps the small number of cases accounts for the fact that there are few peaks in these curves, especially for the boys. The 5 to 8 and 9 to 12 groups for the girls show more of a central tendency.

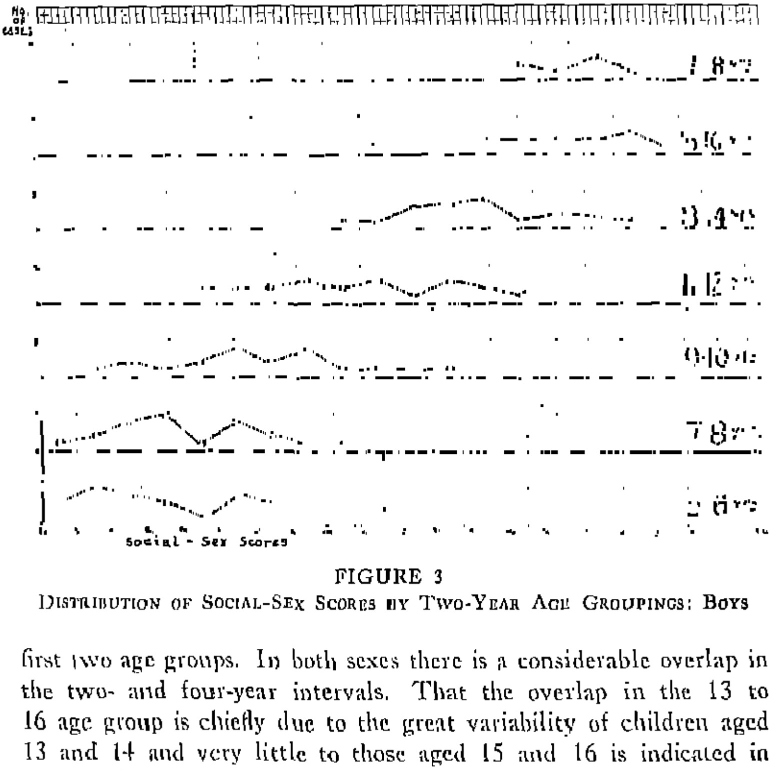


FIGURE 3
DISTRIBUTION OF SOCIAL-SEX SCORES BY TWO-YEAR AGE GROUPINGS: BOYS

Owing to the relatively small number of cases in each age group, the meaning of these distributions should not be stressed, but they do indicate fairly clearly that there is greatest variability in social-sex development in both sexes in the ages from 10 to 14, and that

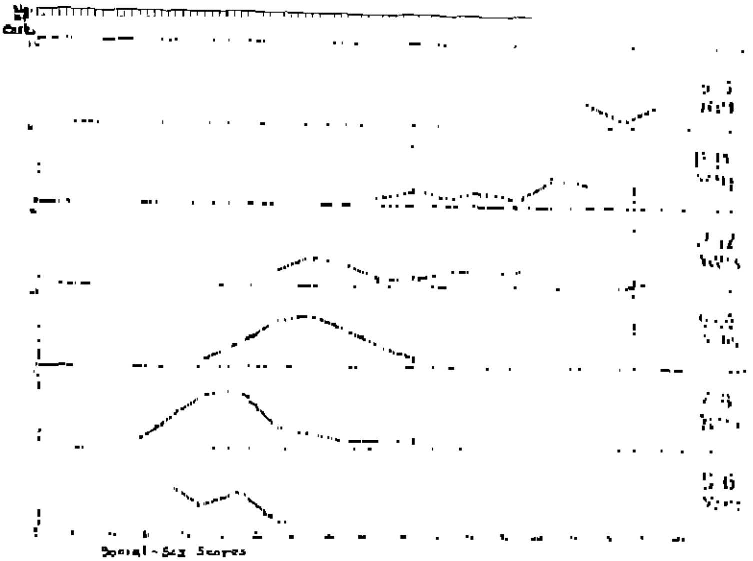


FIGURE 4
DISTRIBUTION OF SOCIAL-SEX SCORES BY TWO-YEAR AGE GROUPINGS: GIRLS

by the time the age of 15 to 16 is reached there is less variability in this respect, the slower ones perhaps catching up with the more precocious. These findings are in line with the present 6-3-3 system of our large city school systems. They may be used as a tentative guide in setting up recreational programs for the two sexes together, where groups are best selected on the basis of social age rather than chronological age.

The percentiles given in Tables 8 and 9, in common with the

TABLE 8
SOCIAL-SEX DEVELOPMENT: SCORES FOR MERRILL-PALMER GIRLS IN TERMS OF PERCENTILE RANK*

Age		Percentile rank														
(years)	*N*	1	3	5	10	20	30	40	50	60	70	80	90	95	97	99
5- 8	38	48	51	55	56	61	64	68	72	75	79	83	91	99	109	126
9-12	40	63	68	72	77	84	89	93	97	101	106	113	127	145	151	161
13-16	25	103	107	111	122	140	154	164	171	173	176	182	189	194	196	199

*The same girl appears twice or oftener in the *N* and scores if she was rated more than once within the age interval.

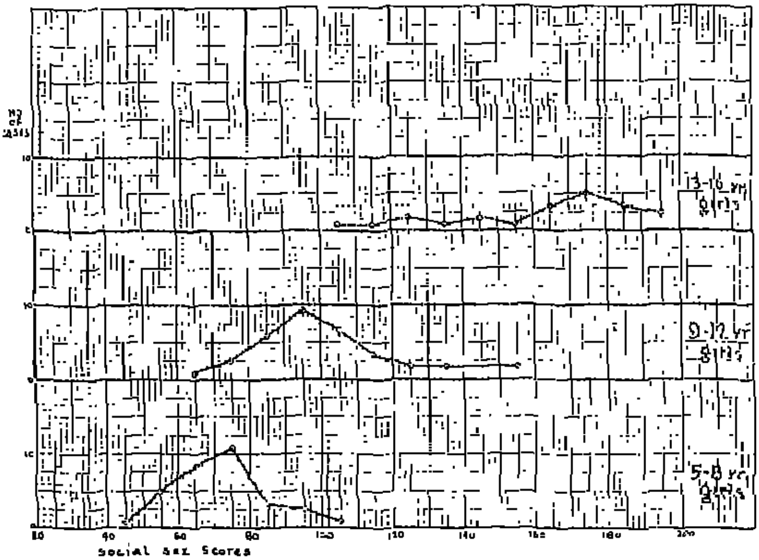


FIGURE 5

Distribution of Social-Sex Scores by Four-Year Age Groupings: Boys

TABLE 9

Social-Sex Development: Scores for Merrill-Palmer Boys in Terms of Percentile Rank*

Age (years)	N	1	3	5	10	20	30	40	50	60	70	80	90	95	97	99
		Percentile rank														
5- 8	44	18	23	25	30	37	41	46	52	59	66	74	83	93	99	108
9-12	38	36	48	53	60	68	76	84	90	96	102	119	129	134	139	154
13-17	30	104	109	113	123	136	145	154	162	169	174	179	184	187	189	193

*The same boy appears twice or oftener in the N and scores if he was rated more than once within the age interval.

other Merrill-Palmer tables (69) based on the average scores of two raters, are offered as tentative only, in view of the small number of cases. The overlap between age intervals, already discussed, makes their use somewhat questionable.

Two methods were tested in making the distributions. In one, the scores of the same child were used two or three times in one four-year age interval if he had records to that extent, just as if the scores represented two or three different children. In the other

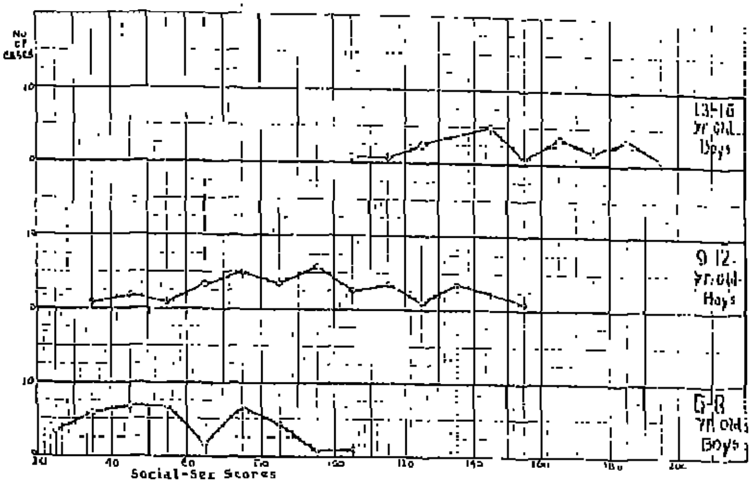


FIGURE 6

Distribution of Social-Sex Scores by Four-Year Age Groupings: Girls

series, such scores were averaged so that the child appeared but once in an age interval. The results as shown in the percentile curves were much the same in both series. However, the norms based upon the second method, and therefore upon a smaller number of cases, were less satisfactory in their lower percentiles and smaller range. Accordingly, Tables 8 and 9 are based upon the first method, where the same child may appear more than once in an age grouping. Further data would doubtless necessitate a revision of the percentiles.

Owing to the large age interval in each percentile distribution, the changing percentile ranks of individual children from year to year are hardly indicative of a change in relative position among their contemporaries. Where a single child's three scores fall within one interval, the average rank ascends more often than not from the first year to the third, because the age interval is too large to exclude the growth factor. When consecutive scores fall in two different percentile distributions, the rank is likely to decrease for the same reason. Until sufficient cases are available to permit making percentile distributions for each year of age, the raw score will probably be better than the percentile rank as a measure of development.

How consistently children are rated in the lower, middle, or upper quartiles is shown for the 34 children rated for three successive years (Table 10).

TABLE 10

Perfect consistency in averaged scores of two raters	12
Consistency two years out of three	20
Lack of consistency throughout	2
Total	34

Thus, 12 of the 34 were given scores in the middle quartiles for three years in succession. Of the eight children given high or low scores in the first year, the detailed tabulation shows that none were placed in that same position for three successive years. Only two were placed in all three positions in the three years. It is possible that some of these cases are to be explained by the fact that the extremes in social-sex development sometimes do not become manifest until the children have attained a certain age, and that in these cases consistency in scores from one age to another cannot be expected. Again, a child may show superior development at one stage as compared with other children, and not appear to be advanced later, when other children are also showing a more advanced kind of behavior.

V. RELATION OF SOCIAL-SEX DEVELOPMENT TO OTHER ASPECTS OF DEVELOPMENT

The correlations shown in Table 11 indicate a very low but positive relationship between skeletal age, mental age, *IQ*, and social-sex development. Obviously quite different aspects of development are being measured.

TABLE 11

	raw *r*	N	*r* age held constant
Skeletal age (Todd's standards)	.714	93	.195
Mental age	.761	99	.118
IQ	—.029	99	.165

VI. USE OF THE SCALE

It has been pointed out that in the use of the social-sex scale percentiles have proved less satisfactory than mean scores. It is therefore recommended that the figures in Table 7, based upon scores falling within the middle two-thirds of the two-year distribution, be used as a tentative standard. However, these scores were obtained from a relatively superior group of children, and possibly, therefore, are applicable only to similar groups.

The procedure is as follows: The observers (one or two) assigned should study the scale and its purpose, then carefully observe the child or children to be measured, in a free play group of boys and girls. If the observer has been working with the child or children for some time not more than four or five periods of observation will be necessary; if not, more time may be required.

As the observer fills in the blank, after leaving the group, he should attempt to base his judgment on remembered examples of the child's usual behavior, and not upon feelings or impressions. The blank should then be scored according to the key, each checked statement being assigned its proper scale value. The average of these values for the whole sheet should then be compared to the scores in Table 7, for the age and sex group to which the child belongs.

If the raters are inexperienced, it is probably better to use the average of the scores given by two raters. In this study, the score of one experienced and reliable rater was found to be preferable to the average of scores given by one good rater and one who was less so. Means of testing the efficiency of raters have already been discussed.

If the child's score falls outside the range shown in Table 7, he may be retarded or advanced in social-sex development. If so, he may be maladjusted in the play group where he has been observed, and some effort should be made toward placing him in a more suitable group. Such a maladjustment is often evident without such a measure, as in the case of boys' and girls' clubs where the majority want to dance and the less-developed minority do not, and spend their time harassing the adult leader and the other children.

It is probable that the scale is most useful as a research instrument, to study consecutive development, and as an administrative instrument, to assemble fairly homogeneous groups for boys' and girls' clubs.

VII. CHARACTERISTICS OF SOCIAL-SEX DEVELOPMENT IN VARIOUS AGE GROUPS, ACCORDING TO THE SCALE

Earlier in the paper the characteristic stages in social-sex development, based upon free observations in boys' and girls' clubs, were described. According to these observations, the general sequence of development was as follows: First, there was a period in the early school years when there was little sex discrimination in play, and a companion of one sex was as acceptable as one of the other; then there was a period, preceding adolescence, when both boys and girls preferred companions of their own sex; and finally, a period of increasing attraction to and companionship with the opposite sex.

How do the findings shown in the scale agree with such a picture? In the following descriptions, the equivalent of the statements in the scale are arranged in the order of their scale values from low to high.

A. Social-Sex Development of the Boy

1. *Youngest Stage* (*Ages 5-8*)

A boy at this stage will play with a group otherwise made up entirely of girls, because he is not yet conscious of sex differences, nor is he embarrassed to be found in such a group. He does not object to having adults of either sex show physical affection for him. He is not yet modest as to posture, gesture, clothing, etc. He does not differentiate games according to sex. He shows no protective habits toward women and girls. When in a game not involving physical skill, he is not inclined to choose his own sex over the other. He fights physically with girls. He is not yet self-conscious or embarrassed by physical contact with girls. He is careless of his personal appearance. Work is work to him, and he does not regard any one kind of work as suitable to boys and another as suitable to girls. The concept of "sissy" is still to be discovered. He is not concerned with girls as attractive creatures. On the whole, he prefers women to girls; at least, women who play with him. He is in a creative period in handicrafts and keeps very busy, and is not tempted to leave his work to play with girls, as he will be later.

2. *Middle Period (Ages 9-14)*

At this stage the boy is found playing pursuit games with girls, such as informal tag games indoors. So much attention he will pay to girls, but in general he shows no interest in what they are doing and even in games not involving physical skill he prefers boys on his "side." When allowed to choose, he always sits next to boys rather than girls. He will not join in a game in which he is the only boy, but must have other boys with him when he plays with girls. Toward the end of the period he becomes sufficiently conscious of sex so that he does not wish to touch girls or show them any attention except under socially approved conditions, such as in games or dancing. If he finds himself in a group of girls, he leaves quickly. Still later in this period he begins a teasing derogatory kind of talk about his friends who have girls, with the intention of "fussing" the boy in question. He is extremely self-conscious and modest about the physical aspects of sex and would not for the world undress or go to the toilet before girls or even women, except where the relationship is parental.

3. *Third Period (14½-17)*

The boy at this stage would not be caught mentioning the fact that a certain girl attracts him, though he is probably beginning to have such feelings, as is evident from greater care for his appearance when he is to meet girls. He is self-conscious and embarrassed when a girl pays him some attention. He has passed the stage when adults are sometimes preferred to boys and girls as playmates in a recreational group, and appears to feel that he has new interests which only his peers can understand. Later in the period, he shys away from girls in a group, but he begins to take some interest and to be less shy when there is only one girl to be considered. He will not tolerate physical affection from adults at this stage; perhaps it implies that he is not yet grown up, and possibly it has some sexual significance to him. He begins to be interested in dancing and has sufficient courage to ask a girl to dance, while before an adult had to persuade him to do so. He begins to enjoy dancing as a part of the total social-sex experience. As yet he has no special choice of girls; almost any girl will do. He begins to wait on women and to be attentive to their needs. He feels that certain kinds of work are suitable for boys, others for girls. Later he is enthusiastic about dancing and prefers it to other group activities, such as singing and games. Later still his interests are sufficiently heterosexual to permit candor about them. He will sit next to his "girl friend" even though he is teased about it. He spends more energy trying to attract the attention of girls and will put down his work to join girls at some play or activity. He imitates adult manners for the first time, to prove his maturity. By now he prefers to sit next to girls and enjoys physical contact with them. He assumes the masculine right to dominate women.

B. SOCIAL-SEX DEVELOPMENT OF THE GIRL

1. *Youngest Stage (Ages 5-8)*

The girl of these ages is perfectly willing to play in groups composed entirely of boys. In choosing sides in games not involving physical skill she has no particular preference for one sex over the other. She is not yet modest about physical matters, will sit in any posture without embarrassment, show her underclothes, and go to the toilet where boys are. She is not embarrassed in a group of boys, as she will be later. She pays no attention if a boy touches her. She likes to have either men or women show physical affection for her. She is careless about her personal appearance, even when with boys. It has never occurred to her that tasks should be allotted on sex lines.

2. *Middle Period (Ages 8½-13½)*

In this period the girl shows no interest in what boys are doing merely because they are boys. She will not stay long in a group of boys if she is the only girl. In choosing sides she is likely to choose girls unless it is a game involving physical skill, when she may choose a boy in the interests of victory. She prefers men to boys. She sits next to girls if given a choice. She will not participate in an activity unless other girls are included. She will invite men to sit next to her, but never boys. She begins to be sufficiently conscious of sex so that she will not deliberately touch boys except under conventional circumstances, as in games or dancing. She classifies games according to sex—boys play this, girls play that.

Later she enters the "whispering period" with her girl contemporaries. She is shyer with a group of boys than with a single boy. If she dances, she prefers to dance with other girls. She would not admit that a certain boy is attractive to her, though she begins to take a covert interest. By this time she is modest about exposing her body and underclothing before boys—probably more so than she will ever be again. She is sufficiently conscious of the sex attraction of clothes to admire the clothing of women and her girl friends. She begins frankly to enjoy dancing.

3. *Third Period* (*Ages 14-16*)

If a boy pays attention to her, she is extremely conscious of it, if one can judge by her embarrassment, interest, and sarcasm. She is beginning to primp and when she is to be with boys spends much time combing her hair, washing, etc. Later she has become sufficiently heterosexual in her interests to sit next to a favorite boy, even though her friends tease her. She has become imbued with the social-sex culture pattern sufficiently to appear to acknowledge the right of boys to dominate over girls. Any occupation she happens to be interested in is dropped if some boy asks her to join him. She has not quite reached the pairing off stage, but accepts the attentions of almost any boy. She definitely classifies work on sex lines—boys move furniture, girls do sewing, etc. Later she is frank in her preference for sitting next to boys. She uses many techniques to attract the attention of boys and is much more interested in what the boys are doing than in what the girls are doing. She assumes many adult airs and manners which may be amusing to an adult, but to her signify her adult status. In contrast to the boy, she indulges freely in physical affection with other girls of her age. She never dances with a girl if boys are present and ask her to dance. By this time she does not accept a boy as a dancing partner unless she deems him sufficiently mature. He must be older than herself, and preferably taller.

Later she seems to enjoy casual physical contacts with boys, as in dancing. She is beginning to have dates. By this time she is overtly heterosexual, and talks to her girl friends about the attractiveness of certain boys.

VIII. SEX DIFFERENCES AND SIMILARITIES IN SOCIAL-SEX DEVELOPMENT

At six, both boys and girls are in a non-sexual or undifferentiated phase of social-sex development. They play with companions of either sex. They are not self-conscious about their bodies and are not embarrassed by physical affection from adults or physical contact with the opposite sex. Boys fight girls and have no special courtesy toward women. They care nothing about their personal appearance. They do not regard any particular work or play as more suitable for one sex than the other. This period lasts until about 8½ years in both sexes. At this time they enter into a unisexual phase characterized by strong attachment to their own sex, and lasting for a shorter period in girls than in boys. Toward the end of the period (11 years) girls are showing self-consciousness at touching boys, except under conventional conditions; boys reach this stage at 12. In this period girls classify games according to sex. At 13 a girl is sufficiently sex-conscious to feel shy in a group of boys. Just previously, at 12, she begins her *incessant whispering with* friends. At 13½ she will not admit that any boy could be attractive to her, but from her behavior one knows she is interested. Boys do not reach this stage until a year later. Sex modesty appears in both sexes at 14.

At 14½ the girl is definitely beginning the heterosexual phase by primping, being enthusiastic about dancing, admiring the clothes of women, and seeming interested in the attentions of boys. The boy of this age is more careful of his appearance and self-conscious about the attention of girls, but he is not interested in dancing until a year later. Boys lose interest in adult affection at 15; girls do not seem to do so.

At 15 a girl is interested in boys, but not in a particular boy; the boy reaches this stage at 15½. Boys show an enjoyment of the physical contact of dancing at 15½; girls at 16. At 15, girls *classify work on sex lines*; boys do so six months later. A definite and open seeking after the other sex begins at 15 for girls, and at 15½ for boys. The acceptance of the respective cultural rôles of men and women is shown in this group at 15 for girls and 15½ for boys.

Girls assume the external manners of an adult at 15½, boys at 16. The patterns are similar for boys and girls, with girls from six months to a year in advance of boys. The sex difference is not as great in this group as some of the literature would lead one to expect; some writers suggest an advance of two years in the girl over the boy.

IX. SEQUENCE OF CERTAIN ASPECTS OF SOCIAL-SEX DEVELOPMENT

A brief analysis of the course of certain special details of social-sex development may be of interest. All descriptions cover the ages from 5 to 16.

A. Attitude Toward Sex of Companions in Play Activities

The boy begins by willingly playing in a group of girls, and is not harassed by the thought that this is unmanly. Team-mates are chosen for skill or friendship *per se*, and not on sex lines. Then comes a time when he will participate in an activity only if other boys are present, and cannot be persuaded to remain in a group in which he is the only boy. He prefers boys to girls in games even when physical skill does not enter in, as in guessing games. Later he has a complete reversal of attitude. Though before he would always choose to sit next to a boy, he will now actually prefer to sit next to a girl, and will do so even under the fire of teasing. He will now remain in a group of girls without the support of other boys for quite a different reason from that prompting him at the earliest stage. In choosing sides for games, he gives girls at least an equal chance because of his interest in them. The girl follows the same course of development in this respect.

B. Attitude Toward Physical Contact

The boy begins by not noticing physical contact with girls, but treats them just as he does boys. He does not object to physical affection from adults. Later he develops self-consciousness about physical contacts with girls, and avoids them except under conventional circumstances, as in games or dancing. Later he appears frankly to enjoy such contacts, in dancing and otherwise. By this time, however, he will not tolerate physical affection from adults, probably because it signifies lack of appreciation of his growing maturity.

The girl follows a somewhat similar course, except that she does not lose her desire for physical affection from adults. In adolescence she likes to show physical affection for other girls and appears to enjoy receiving it from boys.

C. Attitudes Toward Attention Given to and Received from the Other Sex

The boy at first has no observable attitudes toward girls as such. He is not concerned with their desirability or lack of it. In the following, uni-sexual stage, he shows no interest in girls; they are outside his world. Later he shows his consciousness of sex by shyness in a group of girls where he is unsupported by other boys, and is conscious of any special attention paid him by a girl. He begins the period of teasing others about their boy or girl friends. Finally he openly tries to attract the attention of girls, sometimes by amusing means, and when girls are present gives them more time and attention than he gives to boys. He will put aside anything he is doing to join a group of girls who seek him out.

The girl is like the boy in the first two stages. Toward the end of the unisexual period she begins to whisper many secrets to girl friends. She avoids any mention of opinions about the attractiveness of certain boys, though there is evidence that she feels such an attraction. In the beginning of the heterosexual stage she too is self-conscious about attentions paid her by boys, will put aside any activity to join a boy or group of boys, and tries definitely to attract the attention of boys, by both overt and covert means. By this time she talks with other girls about the attractiveness of certain boys.

D. ATTITUDES TOWARD PROPER ACTIVITIES OF BOYS AND GIRLS

At first neither boys nor girls differentiate games and tasks as suitable for one sex or the other. In the middle period they classify games by sex, and, in adolescence, work also.

E. ATTITUDES TOWARD GALLANTRY AND PROTECTIVENESS

The boy begins by fighting with girls, with no feeling that their frailer physique demands concessions. Later his fighting with girls is lighter and not serious, taking the form of light slapping, snowballing, etc. Finally, he waits on girls and shows them a certain amount of protective consideration.

F. ATTITUDES TOWARD RIGHT OF THE MALE TO DOMINATE

In adolescence, along with his protective attitude, the boy assumes in a general way the right of his sex to dominate the other one. The girls seem to accept this assumption. In group meetings, for example, the girls allow the boys to assume a greater amount of leadership than they themselves take.

G. ATTITUDES TOWARD ADULTS

At the beginning of the unisexual stage the boy loses interest in girls of his own age but still shows an interest in women counselors and will include them in a group. In adolescence he definitely prefers girls of his own age to women, but at the same time shows his respect for adulthood by adopting adult manners. At this time he dislikes physical affection from adults.

In the middle period the girl becomes sufficiently conscious of sex to invite a man to sit next to her, though she would not invite a boy of her own age, and she seems to prefer men to boys. In adolescence she admires the clothes of women counselors and imitates their manners, indicating her feeling that she is approaching maturity.

H. ATTITUDES TOWARD DANCING

A sign of adolescence in the boy is a preference for dancing over other social activities. He is not particular about his partner at first, but he has the courage to ask girls to dance and no longer has to be forced to do so by an adult.

The girl begins to dance toward the end of the middle period, but always with other girls; she dislikes to dance with boys. In adolescence she enjoys dancing with almost any boy, and prefers dancing to many other activities. She no longer dances with other girls if she can help it, and prefers boy partners who are older and taller than herself.

I. ATTITUDES TOWARD SEX MODESTY

In the youngest period the boy shows no sex modesty or shyness. Toward the end of the middle period he begins to do so and avoids any bodily exposure before girls. It is at this time that provision must be made for separate toilets and dressingrooms.

The girl at first follows a similar course; by early adolescence she is extremely modest.

J. ATTITUDES TOWARD CLEANLINESS AND APPEARANCE

In the youngest period the boy is entirely carefree about dress and makes no attempt to improve his appearance or make himself attractive. There is no change until the beginning of the heterosexual period, when he becomes conscious and careful of his dress and appearance, presumably to attract the attention of girls. The girl in the youngest stage has no thought of attracting boys by combing her hair, dressing up, washing, etc. By early adolescence she has changed completely, and devotes a great deal of time to primping.

X. COMPARISON OF RESULTS WITH THOSE OF OTHER STUDIES

The findings for boys in the present study agree closely with those of Furfey. At 7½ the Merrill-Palmer boys draw no sex lines and choose girls indiscriminately with boys in games; they also fight physically with girls. At 9½ the unisexual phase begins and continues until up to 14. Furfey places the beginning of this stage at 10, and says it continues until puberty.

Studies from the Catholic University of America on the development of girls also agree fairly closely. In the Merrill-Palmer scale the first statement indicating unisexual interests comes at 8½, but definite preference for play with girls is indicated at 10. Moffitt (46) and Parker (48) found the same to be true. They found heterosexual interests beginning at 14. In the Merrill-Palmer scale suggestions of changes in this direction begin at about 13½. The Merrill-Palmer girls show an overt preference for boys at 15½; the boys show such a preference for girls at 16. The Stanford (65) figures are not comparable since they stop at 14, but at that age show no waning in the unisexual stage except with "gifted" girls.

Hildreth's (27) study shows unisexual interests in 9th grade girls from both public and private schools. The preference of the private-school girls in both the 9th and 12th grades for social dancing compares with the fondness for this activity of the Merrill-Palmer girls of age 14½. Hildreth's subjects in the 9th and 12th grades preferred their contemporaries to adults, a finding that agrees with the result for boys of 15 in the present study; however, this item was discarded in the girls' scale, since it showed no clear differentiation with age.

The finding of Cherkassova (12) of hostility between different-sex partners aged 13 to 16 would probably not be true of the boys and girls in the present study, except as boys of 14 might show hostility. The boys in this Russian study showed greater hostility than the girls, as would be expected in view of the later heterosexual development of boys.

Owing to the failure to limit to definite age periods the terms prepubescence, preadolescence, and adolescence, it is difficult to compare the statements of descriptive writers on these periods with the results of the present study. In general, all these writers agree upon the existence of these three periods and upon the characteristics of them found in the present study. The hostility between the sexes in early adolescence described by some authors might find some verification in the present study if this period were limited to ages 12 to 14. But in girls, at least, such hostility ceases well before the age of 16, according to our findings. Pressey's (52) description follows closely the present findings. We have not found the period of greatest hostility to occur immediately preceding puberty, as so many writers claim, but about two years earlier, unless the teasing between the sexes which is actually an indication of growing interest is looked upon as hostility.

The non-specificity of heterosexual behavior in its early stages, mentioned by many writers, is also shown in the present study.

In our observations on more than 100 children we could find only one child who showed evidence of a "crush," in the sense of an attachment to someone of the same sex. The great hero-worship of adolescence mentioned by many authors was not manifest in our group.

The carefully controlled observations of preschool children made by Parten (50) and Koch (32), showing unisexual preferences, would seem at variance with our findings. It may be that this apparent disharmony can be explained by the fact that at all ages, even through adolescence and adulthood, the sexes tend to be separated in their daily activity, and that had we measured interest by the actual amount of time spent together, or by questionnaires, we might have found boys and girls unisexual in their interests at all ages. In the present study the method was different: we were considering the attitudes of boys and girls to each other when they were brought together in free association in a group.

XI. CASE STUDIES OF SOCIAL-SEX DEVELOPMENT

From the records of the Merrill-Palmer School it is possible to follow the social-sex development of individual children over a period of years. Summaries follow of the relevant records for a girl who was accelerated in social-sex development, though the developmental picture was quite normal, and for a boy who was somewhat retarded in this respect.

A. Case 1

Barbara entered the clubs at the age of 8-10, after a summer at the Merrill-Palmer Camp, and continued to attend until the age of 13-5. She also spent the summer when she was 10-7 at the camp. The records cover this period and represent the findings of the physical growth, mental growth, and educational methods staffs.

Barbara comes from an intellectually and culturally superior home. Both parents are college graduates. The father has had a notable success in his scientific profession, and the mother some success in an artistic endeavor. The socio-economic status is high (Sims Scale score, 34.8). There are two children, the second child being a boy older than Barbara by six years. The house is in a highly restricted residential district. Barbara attended a private girls' school until the depression, when she changed to a public intermediate school. While she has few companions in the neighborhood, and is taken by automobile to her play group, Barbara's play life has been unusually rich, both at home with her brother, parents, her dog, and various equipment, and also in friendships with girls of her own age at school, camp, and clubs, and in the cultural advantages of travel, music and dancing lessons, and attendance at concerts. Her parents have been companionable but have not over-protected her emotionally, and their demanding professions have kept their interests from being too child-centered. Their attitude toward the social-sex development of their children has been almost ideal. They have encouraged a normal, free play life in preadolescence, and when Barbara reached adolescence did not oppose too strongly the external manner of early sophistication which seems to be the present cultural pattern of the urban adolescent. However, Barbara has been expected to conform to a high type of family ideal within this sophisticated pattern, and has so conformed.

She is intellectually in the gifted group. At 9-6 her *IQ* (Stanford-Binet) was 153; at 12-3, 144. According to the Terman group test, her *IQ* at 13-1 was 140, and the grade percentile 97. She is 1½ years accelerated in school. Her school marks are not all high, mainly, her parents believe, because she does little home study.

Her height has been below average and her weight above until the last two measurements at 12-9 and 13-1, when her height reached average; her weight has continued to exceed the average for her height and age by about four pounds. Her parents are of medium height and stocky build. Her health has been excellent except for a food sensitization which she has gradually outgrown. Her food has been well selected and her appetite good. She has slept less than the average child for years, and never seems nervous or overfatigued. She is physically very active. She is a beautiful brunette child, with red cheeks and flashing eyes. All her ratings in physical attractiveness are high and her exceptionally healthy appearance has been noted repeatedly.

In social groups of children Barbara is exceedingly dominating. She has many ideas and prefers to get others to carry them out. She generally has a large following, though not infrequently the followers revolt when they observe that she does not work. However, her originality in planning soon brings them to her beck and call again. Her interests are all outgoing. She demands action of herself and others and is not happy in a contemplative situation. At 10-8 she found it hard to read for an hour at nap time. Her vocabulary level in the mental test was the lowest of all her tests and only a year above her *CA*. On the Marston Scale every rater placed her well on the extrovertive end of the scale.

In a rating of *Attractiveness of Personality* all adults have noted her as possessing chiefly traits on the high end of the scale, but have noted these at the low end at all ages: domineering, inconsiderate of others, quarrelsome, unpopular with other children. In *Independence of Adult Affection or Attention* she is usually marked as above average but shows a tendency to seek the affection and attention of adults through direct and conscious channels. She showed no emotional dependence on adults, however. She is marked high in *Respect for Property Rights* except in a tendency to dislike taking turns. In *Response to Authority* she is noted throughout as at least average, and oftener than not as intelligently coöperative. She sometimes has ideas of her own which she dislikes to interrupt in favor of some adult suggestion. In *Sociability with Other Children* her rating is again high, marred only by her tendency to domineer, and, in her earlier years, to be quarrelsome. An early and continuous friendship with the same girl is noted. Her ratings in *Tendency to Face Reality* are all average, and her failures are in the direction of lack of self-criticism, a liking for the center of the stage, finding it difficult to accept blame, and wishing to be excused from a difficult task. These ratings were made on the *Personality Rating Scales* devised at the Merrill-Palmer School by Roberts and Ball (57).

This general and fairly constant personality rating will serve as a background for a discussion of her social-sex development.

At the age of eight years Barbara was an active, happy, pre-adolescent girl, taking part in all kinds of sports and play, indoors and out, and superior in all of them. She had a close friend, Kitty, with whom she alternately quarreled and made up. She was superior in all creative activities and was far more interested in them than in boys. Activity, and not companionship as such, was the center of her life. She had no sex-consciousness and showed her affection physically for the men counselors at camp just as freely as for the women. At nine, she showed no self-consciousness on being stripped in the weighing and measuring laboratory.

At 10, she continued to succeed in sports, writing, nature study, and handicrafts. She now showed evidence of going through a curious stage girls sometimes experience, when the existence of boys is apparently ignored, though their ideals are aped. At this time Barbara wanted to be a cowboy. She was untidy in her appearance and made fun of girls who primped. She had lost some of her spontaneity of affection for the men counselors, but continued to be very fond of women counselors. She organized a secret "fraternity" of the girls in her cabin, a distinctly unisexual activity. She menstruated during this year.

During her eleventh year, when it was noted that her breasts were well developed, Barbara began to show consciousness of boys; it was noted that she would not participate in a football "huddle" if she had to touch a boy. During this year she spent a large part of her time organizing the *S.S.* club, a secret group within the large Merrill-Palmer club. The author and another counselor were invited to belong, and so many direct observations were possible. The club met in a small, dark space under some basement steps. By dint of much direction from Barbara and work from the members it was made into a clubroom with a library, lights, and a hammock. All the members except the two staff counselors were older girls from the larger club.

and there was much spying on their affairs by the younger girls and the boys. The club had dues, officers, a constitution, a purpose, and existed as a secret, selective, unisexual activity for several months. During the winter the *S.S.* club, with Barbara as leader, organized an elaborate snow fight against the boys. At this time Barbara reached the height of her unisexual period; from this time forward her play shows her developing heterosexual interests.

At the beginning of her twelfth year she began to take an interest in dancing, though she danced badly. She generally danced with girls when she danced at all. She completely planned a musical show, full of semi-sophisticated humor, and dominated the boy who was chairman of the committee. At this time she began to show an interest in boys as such by snatching their handkerchiefs. She did not see why girls and boys should draw sex lines in the duties assigned to them. She still did not feel self-conscious when adults of either sex gave her affectionate pats of approval, but she definitely preferred women counselors to men. She had begun to show an interest in boys, however, for she primped and was interested in boys' activities. She was not embarrassed to find herself in a group of boys and would sit next to a boy if he asked her to do so, but she did not yet openly choose to do so.

In her thirteenth year she was definitely committed to heterosexuality. She preferred dancing to other activities, and with boys more than with girls. She openly preferred certain boys as partners, and would play with her favorite boy friend even when the younger girls teased her. She enjoyed physical contact with boys, at least in dancing and games. She had stopped fighting physically with either sex. She was obviously conscious of and flattered by attention given her by one boy. She tried to get the attention of boys but did not waste much time in getting that of girls. She whispered incessantly to her girl friends, presumably about boys. She imitated adult manners and began to notice and admire the dresses of women counselors. At this time she had lost all her tomboyish desires, and regarded clothes suggestive of boys' clothing as unsuitable. She had also come to regard certain games, such as football, and certain kinds of work, such as carpentry, as unsuitable for her sex. She tried a little to dramatize her sex by appearing gentle, dainty, etc. She was definitely modest, and she began to talk about "dates." However, in spite of this definitely heterosexual picture she still spent more of her time with her girl friends and with women counselors than with boys, probably because of the difference in the physical interests and skills of boys and girls in sports.

B. CASE 2

Alex entered the clubs at 10-4 and still attended at the age of 16. He attended the Merrill-Palmer Camp the summer previous to entering the clubs, and also when he was 11. His parents are both college graduates, and were scholastically brilliant. Their economic status is much lower than Barbara's, for the father has not made a financial success of his profession. The home is in a rundown neighborhood where the neighbors are not the equals of Alex's parents in socio-educational status. The score on the Sims Scale is 30.4, in the 9th decile, which indicates a well-favored home. Alex has a brother and a sister, both some years younger. His play life has been a mixture of play with a rough group of boys in the neighborhood and formal group contacts made for him by his mother, with the Merrill-Palmer clubs, Boy Scouts, choir, and church school. His parents have spent considerable time in taking their children to museums, exhibits, and lectures. Alex has taken music lessons. He goes back and forth to intermediate school on his bicycle and is relatively independent of the family car for transportation. His play life with children of his own age has been definitely restricted, and he has spent more time at home and less outdoors than most boys. His lack of companionship is due partly to a disagreeable personality and partly to the uncongenial boys in his neighborhood. He has spent more time than the average boy in helping with housework and in inventive shop work.

The parents have spent much time with and for their children, and have consulted often with the School. The mother faces the problems of her children realistically in consultation, but in their presence seems to be over-generous in praise. The father has been quick-tempered with the children, but as they grow older spends much time with them. He is somewhat egocentric and unsociable. All three children "show off" before adults by much self-centered talk. Alex is the greatest problem in a difficult family, and has been a difficult member of all social groups for a long period of time. His brother criticizes him to other children, though he is apparently devoted to his young sister.

The records show a decreasing *IQ* for Alex. At 10-1 his *IQ* was 144; at 12-4, 121; at 12-8, 124, according to the Stanford-Binet; at 13-8, 126, according to the Terman Group Test. The test showed him to be accelerated in visual imagery, but not in vocabulary. He attained the 96th percentile for his school grade in this test. He is not accelerated in school, however, and has received only average or below average grades except in shop work, where his grade was *A*. His quotient on the Arthur Scale was 134. He has done little general reading and no home study. He attended the large Detroit public schools until the age of 14, when he was placed in a private boys' school where his parents hoped the competition would move him to greater effort and the after-school sports help in his social adjustment.

His height has been about three inches below the Baldwin-Wood standards over a period of years, and his weight for height and age from five pounds over the average at 10-4 to 29 pounds over at 13-10. These facts suggest a possible endocrine imbalance, yet no such diagnosis has been made. At 15-7 his basal metabolism was found to be normal and he was placed on a reducing diet. He has eaten great and increasing quantities of carbohydrate foods. He is lethargic physically. With his short-waisted, short-necked, stocky body, he presents a rather odd appearance. There are no records of his sexual development.

Alex's group adjustments have apparently been difficult for years. He is generally disliked by both girls and boys, apparently because of his playing for their attention by foolish tricks, tall stories, and boasting; over-domination which fails to meet with any response, a lack of skill in sports, a tendency to retreat to adults by desirable and undesirable devices for getting their attention, and a general lack of appreciation of the need for coöperation in social enterprises. He has been most successful *in social relationships not involving parity, e.g., in helping* adults with work of various sorts, or directing the play of very small girls. He responds to punishment, rather than reason, as a disciplinary measure.

One is not surprised to find his ratings in *Ascendance-Submission* high; in *Sociability*, low. Most of the time he follows such solitary pursuits as handicrafts, though apparently not from choice. In two ratings on the Marston Scale he has been marked well on the extrovertive side. Ratings in *Physical Attractiveness* and *Attractiveness of Personality* repeat and confirm the descriptions already given. Ratings on *Tendency to Face Reality* at all age levels from 10-0 to 13-10 are low, and all raters seem to agree that he lacks ability to criticize himself, that he feels he should have the center of the stage, that he dawdles to avoid a difficult task, seeks to excuse himself, boasts, is a "smart aleck," substitutes talk and day-dreaming for real effort, and retreats to babyish behavior in case of difficulty. At camp one summer he showed several neurotic traits, using imagined illnesses to get attention or to excuse himself from physical effort, and regurgitating food as a protest against being asked to leave a clean plate.

Alex is talented musically. He has absolute pitch, can play the piano by ear, and has a sweet singing voice and perfect rhythm in dancing. But in this area too he lacks sufficient concentration of energy to persist until proficient. Though he is inventive mechanically, his ideas more often than not end in only partial completion of a project, or, worse still, in mere talk.

With this general description of Alex's personality, an attempt will be made to review his social-sex development.

The records for his tenth year show Alex carrying on individual handicrafts and leading a somewhat solitary life, owing to his uncooperative behavior. He apparently had no regard for either unisexual or heterosexual relationships. He had as yet no feeling that one sort of work or occupation was more suitable to girls than to boys, for he indulged readily in such activities as hooking rugs. He paid an equal amount of attention to women and men counselors. He did not follow the other boys in sports.

At 11, he showed much the same picture in his club activities—individual pursuits, but little absorption in them. At camp he made a real effort to lead the group of boys in his cabin, but as usual did not succeed. He was inclined to use salacious language. He showed some tenderness toward his little sister and seemed to enjoy playing with much younger girls, for they were impressed with his big stories and his ideas and liked his domination. According to social-sex ratings at 11-9, he was still willing to sit next to girls if asked to do so, and was willing and unembarrassed to play in a group composed entirely of girls. However, he did spend more time with boys than with girls. He enjoyed the company of women counselors more than that of men, perhaps because he felt less inferior in physical skills with them. He did not object to physical affection from any adult. He seemed to prefer his mother to his father. As yet, he showed no protective manners toward girls and women. He was still extremely careless about his personal appearance, and was not at all concerned about the physical attractiveness of girls. He did not classify work or games according to sex. As will be noted, this description is fairly nearly that of a child in his first period of social-sex development, when Alex should have been well into the second period.

In his twelfth year, Alex continued his solitary handicraft pursuits. There is one observation noting that he apparently felt embarrassment, yet satisfaction, at finding himself at a table alone with a girl. At another time he worked at a painting job with a girl, and told her that he "always got her mixed up with his best girl at school." During the year he took part in social dancing lessons, where his musical talent stood him in good stead, but where he showed no interest in dancing as a social-sex activity. At this time he was still willing to sit next to a girl if requested to do so. He did not show the violent anti-feminine reactions so many boys show at this age. He showed no self-consciousness in physical contacts with girls, and still did not object to physical affection from adults. He was still careless about his personal appearance and unconcerned about the attractiveness of girls. He saw no distinction between work and play for girls and for boys.

During his thirteenth year he was still solitary and did not join in the dancing. At 13-6, however, it was noted that he was getting on better with other boys by joining in a skating party. At 13-10 he was making a bracelet for his "best girl" at school. According to the social-sex observations, he was now definitely in the unisexual phase, seeming to prefer the companionship of boys. No protective attitudes toward women and girls were observed. He refused to dance, was extremely careless about his personal appearance, and still unconcerned about the attractiveness of girls. He did not differentiate between games according to sex. He was not yet modest in posture, gestures, clothing, etc. It may be that his talk of a "girl friend" at this time and previously was wish-fulfillment; or it may indicate that, in his unpopularity with boys, he was obliged to fall back on the companionship of girls. No behavior at any time thus far would lead one to believe that he had any true heterosexual interests.

No consecutive records are available for Alex during his fourteenth and fifteenth years. One student reports that at 14-11 he vied with his brother for dances with a certain girl, but that this was exceptional, and that he usually paid no attention to girls and had to be coaxed to dance. She reports also that he still spent most of his time aimlessly in the crafts room and not on the sports field with the other boys. At this time he made another bracelet for a girl friend, but it was never known whether he gave these bracelets away. A staff member reports that in his fifteenth year Alex was still apparently in the unisexual phase and when possible refused to dance.

C. Comparison of the Two Cases

It is interesting to speculate about Barbara's acceleration and Alex's retardation in social-sex development. Both were of superior intelligence and had well-educated parents with attitudes conducive to normal social-sex development in their children. Both were talented in the arts and crafts, imaginative and creative, both domineering to a degree, both extrovertive, and both had had experience in co-educational public schools and one-sex private schools. The girl had a premature sexual maturing, the boy presumably a retarded one. The girl was well-adjusted in a social group and, except when she was too domineering, carried the group with her. The boy was never popular and was heartily disliked by most children. He faced reality less well than the average, while the girl rarely retreated from it. Was it the social inadequacy of the boy which made him slow to develop an interest in girls, or does sexual immaturity explain his social-sex retardation? Did the girl develop rapidly in social-sex relations because of her early sexual maturation, or because she was well adjusted? Todd found sexual maturity and skeletal age to be related, while in this study social-sex age has a correlation of but .195 with skeletal age with chronological age constant. It is probable that sexual maturity makes a definite, though small, contribution to social-sex maturity.

Alex's retarded social-sex development may be explained partly by the fact that his interests are not those of other boys. Since in our culture interests are supposedly sex-linked, a boy who followed the more feminine type of play would be cast off from his own sex group, and confused sex feelings might follow, according to the reasoning in Mead's *Sex and Temperament*. However, in this case it would appear that Alex's egocentricity was a more obvious cause of his unpopularity with boys than was his lack of athletic interests.

XII. SUMMARY AND CONCLUSIONS

This study attempts to define, describe, and measure the social-sex segment of child development. Social-sex development is defined as the child's social relationships with the opposite sex, leading toward heterosexual adjustment in adolescence. These relationships seem to certain observers to fall into sufficiently well-defined patterns at different stages in childhood to permit using these patterns as norms. Whether the pattern is biologically or socially determined is difficult to say, but there is sufficient evidence from the anthropologists to lead one to believe that it is partly socially determined. Where a culture has demanded a separation of the sexes during preadolescence or adolescence, in work and play, the social-sex behavior is obviously of a more prescribed and conventional pattern than in a culture where more freedom is allowed. The separation along sex lines before puberty in cultures where free play is allowed between boys and girls seems to be common, and one might assume that the separation is due to the diversity of play interests of the sexes. However, since this very diversity is probably in part culturally determined, and in part organically conditioned, the fairest conclusion seems to be that the social-sex pattern has roots in both biological and social determinants. The pattern described in this study, resulting from observations made in the United States, among boys and girls gathered in free play groups, would not therefore apply in detail in other cultures.

Such a study is not the less important, however, because it deals with a pattern that is chiefly culturally determined and hence is limited in its application. The difficulties in adjustment of a child who for some reason differs notably from the norm, whether biologically or culturally determined, are patent. As Margaret Mead points out in *Sex and Temperament*, it makes no difference to the unaggressive boy that aggression is no more sex linked than blue eyes; he suffers deeply because society demands that boys be aggressive, while he is not.

The practicality of such a study to educators is evident when one considers the claims which are made for the coeducational school as an instrument in developing normal heterosexuality, or the unisexual play group, camp, or school as more pleasurable to the children enrolled.

This study does seem to support the claim that at least during one stage in childhood the sexes prefer to be separated. However, when the separation of the two sexes takes place voluntarily, with the two groups still in propinquity, it appears to the author, though it is not proved experimentally, that the normal drawing together in adolescence is accomplished with less awkwardness than when an overt break is made earlier. From a reading of the anthropologic literature one infers also that when the sexes are kept separated in adolescence, there is more homosexuality than in cultures which allow more freedom.

In general, the social-sex pattern develops as follows: There is at first an undifferentiated social relationship with the opposite sex until about the age of eight years, then a rising preference for children of the same sex, until puberty, when heterosexual feelings begin gradually to develop.

Method. The present study began with observations of recreational groups with a membership of 112 former nursery-school children, 53 girls and 59 boys, aged 5 to 17, above average in intelligence and socio-economic status, where the program was free enough to allow the boys and girls to form activity groups of whatever sex composition was preferred. These observations, recorded in narrative form, were then made into a series of short statements descriptive of different aspects of social-sex behavior. Paired observers, not always together, then checked the appropriate statements for each child. This process was repeated annually for three successive years. At the end of this period an attempt was made to scale the statements by two methods: (*a*) by age, and (*b*) by a coarse scoring method related to age. The second method was found to be as effective as the first, and simpler to use, and was therefore adopted. All statements not showing a clear age differentiation were discarded. The number was further reduced by eliminating repetitious statements and also by reducing an overabundance of items in the older age ranges by means of a finer statistical and judicial discrimination.

Results. Separate scales for boys and girls were produced. The final scales do not have equidistant steps, but there is no conspicuous irregularity.

A study of the reliability of the scale shows a degree of agreement between the results of paired ratings with age held constant, comparing favorably with the results of other studies. A detailed comparison of the ratings made by individual raters indicates that where, of one pair of raters, one is found to have a constant bias toward higher ratings and the other toward lower, an average of their two scores may offer the best means of arriving at the true score. However, in the present study scores depending upon the judgment of one experienced and reliable observer were found to be the most satisfactory.

Since the method used in this study largely depends upon the accuracy of adults' observations of children, a careful study was made of the reliability of raters and ways of testing it. In summary, the findings were as follows: (*a*) Agreement between raters is somewhat dependent upon educating raters to the meaning of the concepts implicit in the scale. (*b*) It is easier to reach agreement in the rating of adolescents than in rating children aged about 9 to 12, and easier, in turn, to agree upon the latter than upon children aged 5 to 8; accordingly, raters should be instructed especially upon the phases of social-sex development in younger children. (*c*) The efficiency of a rater in using the scale may be judged upon the following basis: Not more than 10 per cent of the second year's ratings of a child should fall more than 10 scale points below those of the first year; not more than 20 per cent should be lower to *any* degree; between 40 and 45 per cent should increase more than 10 score points; at least 70 per cent should increase by some amount, large or small.

Tables of mean scores for boys from 5 to 18 and girls from 5 to 16, by two- and four-year intervals, are given. Tentative percentiles are also given, but are less satisfactory for use with the scale.

The distribution of scores indicates that: (*a*) The scores for girls are more advanced than those for boys at all ages. (*b*) The greatest variability in social-sex development in both sexes occurs in the age group from 10 to 14.

Directions for the use of the scale are given. In terms of the statements of the scale, descriptions are given of the social-sex development of the boy and the girl in the three characteristic periods—youngest, middle, and older—the corresponding ages varying somewhat for boys and for girls.

The sex differences and similarities in social-sex development and the sequence of certain aspects of this development are discussed. The findings are compared with those of other studies in this area of development. In conclusion, there are two illustrative case studies: one of a girl accelerated, one of a boy retarded, in social-sex development.

REFERENCES

1. Averill, L. A. Adolescence: A Study in the Teen Years. Boston: Houghton Mifflin, 1936.
2. Blanchard, Phyllis. The Adolescent Girl. New York: Dodd, Mead, 1926.
3. Bolton, F. E. Adolescent Education. New York: Macmillan, 1931.
4. Bonner, A. I. Emotional Problems of Adolescents. Chicago: Univ. Chicago Press, 1930.
5. Boorman, W. R. Developing Personality in Boys. New York: Macmillan, 1929.
6. Brooks, F. D. The Psychology of Adolescence. Boston: Houghton Mifflin, 1929.
7. Bühler, Charlotte. The social behavior of children. *Handbook of Child Psychology*. Worcester, Mass.: Clark Univ. Press, 1933.
8. ———. Three Generations of Youth as Seen in Their Diaries. Jena: Fischer, 1934. (From *Psychol. Abstracts*, VIII, 2329.)
9. Burks, B., Jensen, D. W., & Terman, L. M. Genetic Studies of Genius, Vol. III. The Promise of Youth. Stanford Univ.: Stanford Univ. Press, 1930.
10. Busch, H. M. Leadership in Group Work. New York: Association Press, 1934.
11. Campbell, Elise H., & Tyler, Dorothy. A recreational program for children from five to fifteen. *Childhood Educ.*, 1934, **11**, 22-28.
12. Cherkassova, A. On the question of studying experimentally the collective creativity of the child. *Dyetskii kollektiv i rebenok*, 1926, 138-196. (Quoted from Murphy and Newcomb, *Experimental Social Psychology*. New York: Harper's, 1937.)
13. Cole, L. Psychology of Adolescence. New York: Farrar & Rinehart, 1936.
14. Dimock, H. S. A research in adolescence: The social world of the adolescent. *Child Devel.*, 1935, **6**, 177-195.
15. Dunlap, K. Civilized Life. Baltimore: Williams & Wilkins, 1934.
16. Elliott, G. L. Understanding the Adolescent Girl. New York: Holt, 1931.
17. Elliott, G. L., & Bone, H. The Sex Life of Youth. New York: Association Press, 1929.
18. Ellis, H. Studies in the Psychology of Sex. Philadelphia: Davis, 1929.
19. Flory, C. D. Sex differences in skeletal development. *Child Devel.*, 1935, **6**, 205-212.
20. Furfey, P. H. A revised scale for measuring developmental age in boys. *Child Devel.*, 1931, **2**, 102-114.
21. ———. The Gang Age: A Study of the Preadolescent Boy and His Recreational Needs. New York: Macmillan, 1926.
22. ———. The Growing Boy: Case Studies of Developmental Age. New York: Macmillan, 1930.
23. Fuxloch, K. Das Sociologische im Spiel des Kindes. (Social factors in the play of children.) *Beihefte z. Zsch. f. angew. Psychol.*, 1930, No. 53. Pp. 96.

24. Goodenough, F. L. Developmental Psychology. New York: Appleton-Century, 1934.

25. Green, E. H. Friendships and quarrels among preschool children. *Child Devel.*, 1933, **4**, 237-252.

26. Hagman, Elizabeth P. The Companionships of Preschool Children. Iowa City: Univ. Iowa Stud. Child Wel., 1933, **7**, No. 4. Pp. 69.

27. Hildreth, Gertrud. Adolescent interests and abilities. *J. Genet. Psychol.*, 1933, **43**, 77-78.

28. Hollingworth, Leta S. The adolescent child. *Handbook of Child Psychology.* Worcester, Mass.: Clark Univ. Press, 1933. Pp. 885-892.

29. ———. The Psychology of the Adolescent. New York: Appleton-Century, 1928.

30. Hurlock, E. B., & Klein, E. R. Adolescent "crushes." *Child Devel.*, 1934, **5**, 63-80.

31. Isaacs, Susan. Social Development in Young Children. New York: Harcourt, Brace, 1933.

32. Koch, H. L. Popularity in preschool children: Some related factors and a technique for its measurement. *Child. Devel.*, 1933, **4**, 164-175.

33. Lehman, H. C., & Witty, P. A. The Psychology of Play Activities. New York: Barnes, 1927.

34. Lynd, R. S., & Lynd, Helen M. Middletown. New York: Harcourt, Brace, 1929.

35. Macauley, E. Some notes on the attitude of children to dress. *Brit. J. Psychol.*, 1929, **9**, 150-7.

36. McCarthy, R. C. Training the Adolescent. New York: Bruce, 1934.

37. Maller, J. B. Coöperation and Competition: An Experimental Study in Motivation. New York: Teachers College, Columbia Univ., Contributions to Education, No. 384, 1929.

38. Mandelbaum, May. Batciga education. Hanover Conference Seminar on Human Relations, 1934.

39. Mead, Margaret. Coming of Age in Samoa. New York: Morrow, 1928.

40. ———. Arapesh Education; Education among the Ba Thonga; Education among the Cheyenne Indians; Education among the Manus of the Admiralty Islands; Education in Samoa; Education and the Community in Samoa. Hanover Conference: Seminar in Human Relations, 1934.

41. ———. Growing up in New Guinea. New York: Morrow, 1930.

42. ———. Sex and Temperament in Three Primitive Societies. New York: Morrow, 1935.

43. ———. The primitive child. *Handbook of Child Psychology.* Worcester, Mass.: Clark Univ. Press, 1933.

44. Miles, Catherine C. Sex in social psychology. *Handbook of Social Psychology.* Worcester, Mass.: Clark Univ. Press, 1935.

45. Mira, E. Psicopedagogia de la sociabilitat. (Educational psychology and socialization.) *Rev. Psicol. Pedag.*, 1933, **1**, 354-368. (From *Psychol. Abstracts*, IX, 4286.)

46. Moffitt, Constance. The Girl Grows Up: A Study of the Development of Personality in Girls Ten and Twelve Years of Age. Mas-

ter's thesis. Washington, D. C.: Catholic University of America, 1930.

47. Morgan, J. J. B. Psychology of the Unadjusted School Child. New York: Macmillan, 1926.

48. Parker, Norma. The Girl Grows Up: A Study of the Development of Personality in Girls Fourteen and Sixteen Years of Age. Master's thesis. Washington, D. C.: Catholic University of America, 1930.

49. Parshley, H. M. The Science of Human Reproduction. New York: Norton, 1933.

50. Parten, M. Social play among preschool children. *J. Abn. & Soc. Psychol.*, 1933-4, **28**, 136-147.

51. Pechstein, A., & McGregor, A. L. Psychology of the Junior High School Pupil. Boston: Houghton Mifflin, 1924.

52. Pressey, S. L. Psychology and the New Education. New York: Harper, 1933.

53. Rauth, J. E., & Furfey, P. H. The maturational factor in adolescent conduct. *Child Devel.*, 1933, **4**, 90-93.

54. Read, C. S. The Struggles of the Male Adolescent. London: Allen & Unwin, 1928.

55. Richmond, Winifred V. An Introduction to Sex Education. New York: Farrar & Rinehart, 1934.

56. ———. The Adolescent Boy. New York: Farrar & Rinehart, 1933.

57. ———. The Adolescent Girl. New York: Macmillan, 1926.

58. Roberts, Katherine E., & Ball, Rachel S. A study of personality in young children by means of a series of rating scales. *J. Genet. Psychol.*, 1938, **52**, 79-149.

59. Schwab, S. I., & Veeder, B. S. The Adolescent: His Conflicts and Escapes. New York: Appleton-Century, 1929.

60. Shuttleworth, F. K. The Adolescent Period: A Graphic and Pictorial Atlas. *Monog. Soc. Res. Child Devel.*, 1938, **3**. Pp. 246.

61. Stone, C. P., & Barker, R. G. On the relationships between menarcheal age and certain aspects of adolescent attitudes and interests, intelligence, and physique. Stanford Univ.: unpublished paper, presented in conference.

62. Strain, F. B. New Patterns in Sex Teaching. New York: Appleton-Century, 1934.

63. Stutsman, Rachel. What of Youth Today? Detroit: Y. M. C. A., 1935. Pp. 232.

64. Sullivan, Sister Celestine. A Scale for Measuring Development Age in Girls. Catholic Univ. of America: Stud. Psychol. & Psychiat., 1934, **4**. Pp. 65.

65. Terman, L. M. Genetic Studies of Genius. Vol. I. Mental and Physical Traits of a Thousand Gifted Children. Stanford Univ.: Stanford Univ. Press, 1925.

66. Thom, D. A. Guiding the Adolescent. Washington: Government Printing Office, 1933.

67. Westbrook, C. H., Lai, D. G., & Hsiao, S. D. Some physical aspects of adolescence in Chinese students. *Chinese Med. J.*, 1934, **48**. (From *Psychol. Abstracts*, VIII, 4336.)

68. Unruh, Isabelle. A Preliminary Investigation of a Method of Constructing a Genetic Scale for Measuring the Social-Sex Development of the Child. A thesis presented for the M.A. degree, School of Education, University of Michigan, August, 1935.

69. Wilson, C. A., Sweeny, M. E., Stutsman, R., Chesire, L. E., & Hatt, E. The Merrill-Palmer Standards of Physical and Mental Growth. Detroit: The Merrill-Palmer School, 1930.

70. White House Conference: Social Hygiene in the Schools. New York: Century, 1932.

71. Williams, F. E. Adolescence: Studies in Mental Hygiene. New York: Farrar & Rinehart, 1934.

72. Zachry, C. B. *Social adjustment and sex education. J. Nat. Educ. Assoc.*, 1932, **21**, 5-6.

Genetic Psychology Monographs (*continued*)

VOLUME 17—January-December, 1935

1. Organization of Behavior in the Albino Rat—R. L. Thorndike
2. Brightness Discrimination in the Rhesus Monkey—M. P. Crawford
3. The Limits of Learning Ability in Cebus Monkeys—A. M. Koch
4. Nature-nurture and Intelligence—A. M. Leahy
5. On Intelligence of Epileptic Children—E. B. Sullivan and L. Gahagan
6. A Study of the Play of Preschool Children by an Untrained Observer—D. L. Cockrell

VOLUME 18—January-December, 1936

1. Sex Differences in Variational Tendency—Q. McNemar and L. M. Terman
2. The Process of Learning to Dress among Nursery-School Children—C. B. Key, M. R. White, A. B. Heiney, M. P. Honzik, and D. Erwin
3. A Study of the Present Social Status of a Group of Adults Who, When They Were in the Elementary Schools, Were Classified as Mentally Deficient—Warren Baller
4. The Influence of Specific Experience upon Mental Organization—Anne Anastasi

5 & 6. Studies in Aggressiveness—Lauretta Bender, Sylvan Keiser, and Paul Schilder

VOLUME 19—January-December, 1937

1. Psychological Bases of Self-Mutilation—Casimir Dabrowski
   Masculinity, Temperament, and Secondary Sex Characteristics: A Study of the Relationship between Psychological and Physical Measures of Masculinity—Howard Gilkinson
2. A Psychological Study of Forty Unmarried Mothers—Ruth D. Nottingham
   Behavior Problems in the Children of Psychotic and Criminal Parents—Lauretta Bender
3. Domination and Integration in the Social Behavior of Young Children in an Experimental Play Situation—Harold H. Anderson
4. The Sequential Patterning of Prone Progression in the Human Infant—Louise Bates Ames

VOLUME 20—January-December, 1938

1. The Relationship between Characteristics of Personality and Physique in Adolescents—P. S. de Q. Cabot
2. Behavior Problems of Elementary School Children: A Descriptive and Comparative Study—Isabel Young-Masten
   Graphic Representation of a Man by Four-Year-Old Children in Nine Prescribed Drawing Situations—Pearl Fairfull Gridley
3. Differences between Two Groups of Adult Criminals—Ruth Sherman Tolman
4. A Comparative Study by Means of the Rorschach Method of Personality Development in Twenty Pairs of Identical Twins—Evelyn Troup
   Individual Differences in the Facial Expressive Behavior of Preschool Children: A Study by the Time-Sampling Method—Carla Swan

VOLUME 21—January-December, 1939

1. An Experimental Analysis of "Level of Aspiration"—Rosalind Gould
2. Some Light on the Problem of Bilingualism as Found from a Study of the Progress in Mastery of English among Preschool Children of Non-American Ancestry in Hawaii—Madorah E. Smith
3. Domination and Social Integration in the Behavior of Kindergarten Children and Teachers—Harold H. Anderson
   The Capacity of the Rhesus and Cebus Monkey and the Gibbon to Acquire Differential Response to Complex Visual Stimuli—William Egleston Galt
4. The Social-Sex Development of Children—Elise Hatt Campbell

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Preparation of Manuscripts for The Journal Press

A. Divisions

I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

A. The Next Subdivision Title is Printed in Caps and Small Caps Centered on the Page

1. *Then Italics, with Principal Words, Upper and Lower Case, Centered on the Page*

a. Then italics, upper and lower case, 1-em run-in side head.

(1). Then italics, upper and lower case, 2-em run-in side head.

(a). Then italics, upper and lower case, 3-em run-in side head.

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

B. Tables

Tables should always be typed on separate sheets of paper, numbered, and referred to in the text by number (Table 2). Tables should be so constructed that vertical lines are unnecessary.

C. Cuts

Drawings should be made on white drawing paper or on yellow or blue lined graph paper. Lettering should be large enough to be visible after photographic reduction. Lines should be heavy enough to be visible after photographic reduction. Drawings should be numbered, and referred to in the text by number (Figure 2). The legends should be typed on separate sheets of paper. In making drawings only black India ink should be used. Never write instructions for the engraver, as your words may mean the opposite of what you intend.

D. References

References should be segregated at the end of the paper, should be arranged in alphabetical order by author, numbered and referred to in the text by number (2). The proper form for a book reference is as follows:

2. Doe, J. The Preparation of Manuscripts. New York: Holt, 1938. Pp. 400.

The proper form for a journal reference is as follows:

2. Doe, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1938, 19, 550-575.

If in the text it is desirable to refer to a page, thus (2, p. 45).

E. General Cautions

1. Use heavy typewriter paper, double-space the lines, and leave margins for editorial work.

2. Use as few footnotes as possible, but number such as you use, thus.[2]
[2]Footnote.

3. Retype any page on which written corrections have been made. Failure to do this will invite grief.

4. The Editor reads every manuscript carefully for meaning, and may modify punctuation in order to make meaning more clear. If you disagree with this on your page-proof, please initial your objection. The Editor will always yield on such matters to the author, but may not yield to the mechanical objections of a secretary.